SALIENCE OF ERRORS TO LEARNERS OF FRENCH: TOWARD A CHARACTERIZATION OF SALIENCE IN SECOND LANGUAGE ACQUISITION

By

KRISTA HELENE THOREN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1993

Dedicated to the memory of my parents

Zandra Sloan Thoren	1935-1986
Victor Eugene Thoren	1935-1991

ACKNOWLEDGMENTS

I wish to thank students and instructors at the University of Florida and the University of Texas.

I am especially grateful to David Birdsong for his guidance, and to Andrea Tyler, Jean Casagrande, Ira Fischler, and Jeff Farrar for comments on drafts of the dissertation. I also wish to thank Randy Diehl and Cheney Crow for help in acquiring equipment and subjects, respectively, at U.T. Thanks also go to David Johns for writing the computer software, to John Bro for helpful discussions over the years, and to Chuck Kincaid, for help with statistics.

My family has provided much support. I thank my wonderful sisters Vikki, Lori and Paula, and my brothers-in-law Dan and Adrian.

Thanks go also to the following friends: Aaron, Alice, Cathy, Cynthia, Diana, Diane, Gwena, Ivy, Joy, Marty, Mary Lou, Mira, Nina, Patti, and Wren.

Last, but definitely not least, I wish to thank Carl Turner for his patience with me during the writing of my dissertation. I thank him for helping me with numerous aspects, including (but not limited to) software, statistics, and figure drawings. I am grateful for his emotional support.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS.....iii

ABSTRACT.....vi

CHAPTERS

1 INTRODUCTION.....1

1.1 Overview.....1
1.2 Definitions of Salience.....2
1.3 The Relevance of Salience to SLA.....12
1.4 Statement of the Problem.....22
1.5 Notes.....25

2 EXAMINATION OF THE RESEARCH QUESTIONS.....26

2.1 Is Salience a Unitary Concept?.....26
2.2 Is Salience a Stable Concept?.....32
2.3 Is Salience of Types of Errors a Meaningful Notion for SLA?.....48
2.4 Review of the Research Hypotheses.....55
2.5 Notes.....57

3 THE STUDIES.....59

3.1 Tasks and Dependent Measures Used in the Experiments.....59
3.2 Restatement of the Research Hypotheses.....70
3.3 Experiment 1.....72
3.4 Experiment 2.....101
3.5 Notes.....110

4 GENERAL DISCUSSION.....111

4.1 Review and Discussion of the Major Findings...111
4.2 Frameworks to which Salience is Relevant.....115
4.3 Summary.....131
4.4 Notes.....131

5 CONCLUSION.....133

5.1 Limitations of the Research.....133
5.2 Problems in Interpretation.....134
5.3 Pedagogical Implications.....136
5.4 Directions for Further Research.....137
5.5 Closing Remarks.....137

REFERENCES.....141
APPENDIX 1: QUESTIONNAIRE.....153
APPENDIX 2: STIMULI.....154
APPENDIX 3: SAMPLE GEFT ITEM.....155
APPENDIX 4: EXPERIMENT 2 RAW DATA.....156
BIOGRAPHICAL SKETCH.....157

Abstract of Dissertation Presented to the Graduate School of
the University of Florida
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

SALIENCE OF ERRORS TO LEARNERS OF FRENCH: TOWARD A CHARACTERIZATION OF SALIENCE IN SECOND LANGUAGE ACQUISITION

By

Krista Helene Thoren

May, 1993

Chairperson: Dr. David Birdsong
Major Department: Program in Linguistics

Second language acquisition (SLA) research has focused little attention on learner reactions to errors in the target language. This dissertation investigates the salience of errors to second language learners by analyzing responses gathered from three types of tasks: Error detection (timed), judgment of error importance, and judgment of error prominence.

Investigators appear to assume not only a general consensus about what salience is, but also that their own use of the term is consonant with that consensus. The notion of salience, as currently employed both in the literature in general and in the SLA literature, is imprecise and results in confusion when one attempts to analyze its effects on performance.

The current study explores the idea that three qualities, importance, prominence, and detectability, are all components of salience. By investigating how well these three proposed components correlate with each other regarding second language learners' reactions to errors in the target language (French), we address the question of whether or not salience is a unitary concept.

Additional assumptions often made in the second language acquisition literature are that salience values do not change as a learner gains expertise; i.e. that salience is stable, regardless of the learner's proficiency level. This assumption leads to a further assumption that it makes sense to speak of an item's inherent salience, independent of learner expertise level and the interpretation given to the term "salience." The current study compares responses made by subjects of three different proficiency levels (beginning, intermediate, and near native) in an effort to determine whether salience is a stable concept.

Issues of whether salience is unitary and whether it is stable are explored along linguistic dimensions by the use of errors involving three different types of constructions: preposition stranding, double object dative constructions, and relativized resumptive pronoun constructions.

While correlations were found between reaction time and prominence and between prominence and importance, none were found between reaction time and importance. Intergroup differences were found for reaction time and prominence

judgments, and for type of error. Results indicate that salience may be neither unitary nor stable.

CHAPTER 1
INTRODUCTION

1.1. Overview

A major motivation for this dissertation is the conviction that salience is a topic which deserves more attention than it has received, either in the general literature or in the second language acquisition literature. More specifically, little reliable information is available about what linguistic constructions, or errors within those constructions, are salient to second language learners. Accordingly, this dissertation explores salience within the domain of second language learning.

Chapter 1 first examines how salience has been defined, in both the general literature and in second language acquisition (SLA) literature. Next, the relevance of salience to specific issues in SLA theory (e.g. definitions of intake, the role of negative evidence, etc.) is discussed. The research questions to be explored in the current study are delineated in section 1.4.

Chapter 2 examines each of the research questions in detail. Relevant topics (e.g. the novice/expert distinction, metalinguistic performance, etc.) are discussed, and the current research hypotheses are stated.

Presentation of the experimental studies is found in chapter 3.

Chapter 4 contains a general discussion of the results from the studies, and an evaluation of the role of salience within several current theoretical frameworks (e.g. Bialystok's Metacognitive Framework).

Finally, chapter 5 concludes the dissertation with discussions of the limitations of the research, pedagogical implications of the findings, and directions for further research.

The first task, then, is to lay a foundation for a consideration of salience. Toward that end, we now examine how other researchers have defined salience.

1.2. Definitions of Salience

Mention of salience is widespread in the general literature of both psychology and linguistics. Salience has been found to be a factor in perception (Odom and Lemond, 1975; Cunningham and Odom, 1978; Odom, 1982), in memory (Chafe, 1977), in performance of metalinguistic tasks (Galambos and Hakuta, 1988) and in sentence production and processing (MacWhinney, 1977; Osgood and Bock, 1977; Levelt, 1989). In addition, salience has been identified as a determinant of what is noticed both in L1 acquisition (Peters, 1985) and in second language acquisition (Gass, 1980; Chaudron, 1983a; VanPatten, 1984; Robb et al., 1986; Bardovi-Harlig, 1987). Some SLA researchers have proposed

salience as a factor in determining acquisition orders (Gass, 1980; Bardovi-Harlig, 1987). Birdsong and Kassen (1988) noted a possible influence of error salience on subjects' performance in judging the seriousness of errors.

Although salience of a given item is often cited as a factor influencing human behavior on a variety of tasks, discussions of salience often do not define salience at all. Instead, they focus on factors which affect salience. Levelt (1989) discusses discourse factors, such as whether an entity has been topicalized and whether it was recently mentioned. Either of these situations might increase salience. Fillmore (1977) discusses factors relating to what can be called "human interest": namely, factors such as humanness (of more interest to humans than nonhumanness), change of state (more interesting than a lack of change) and definiteness. These factors, he says, affect an item's salience. Naturalness has been mentioned by Hatch (1983), who states that free morphemes are easier for learners to incorporate in their interlanguage than are bound morphemes, because the former are stressed and therefore perceptually prominent. She equates perceptual prominence with salience. Chafe (1977) and Clark and Clark (1977) state that unexpectedness of a semantic item contributes to its salience. They consider salience to be information value. Osgood and Bock (1977) determined that (visual) vividness increases salience. Other perceptual factors such as size, distance, and figure vs. ground (the former is more salient)

are well documented as factors affecting salience. In addition, increased imageability may make a noun more salient (Paivio et al., 1968; Levelt et al., 1977). Concreteness appears to be the major, though not the sole, determinant of a noun's imageability. For example, "table" is a more concrete noun than "love," and thus has higher imageability. Strickland and Ayoun (1989) and VanPatten (1984) suggest that content words are more salient to L2 learners than are function words. Reber (1980) found that an artificial grammar is made more salient by the way in which it is presented (a set of three strings, aligned vertically), and that a presentation matrix of similar strings is more salient than a presentation of dissimilar ones. He also found (1976) that violations in artificial grammars are most easily detected when the deviant symbol is either at the beginning or at the end of the string. MacWhinney (1977) suggests that agency helps make a grammatical element attentionally salient.

It is clear, then, that many claims have been made about salience; much attention has been given to factors which might affect it. Far less attention has been given to a determination of what salience itself is. In other words, discussions of what makes an item more salient often exist in the absence of any discussion of what is meant by the term "salient"; salience itself is only rarely defined. This dissertation makes use of those definitions which do exist as a first step in an effort to characterize salience

in second language acquisition. We first examine definitions which appear in the general literature.

Some terms employed as synonyms of salience in the general literature are the following: 1) importance, 2) prominence, and 3) perceptibility or detectability. Levelt (1989) seems to equate salience and importance of a given message in his discussion of factors influencing salience. Peters (1985) discusses salience or importance of a given stretch of speech in terms of what is perceived by the child and therefore extracted from the speech stream. She discusses what certain features of speech (e.g. meaning, suprasegmentals, etc.) do to aid the child L1 learner.

Some researchers appear to equate salience with prominence. Levelt, in addition to employing the term "importance," as mentioned above, also refers to salience as prominence when he discusses factors influencing salience. Other researchers, even in contexts where prominence seems to be the definition employed of salience, do not use "prominence," or any word other than salience, in the discussion. For example, Robb, Ross, and Shortreed (1986) discuss four methods of providing corrective feedback on written English as a Foreign Language (EFL) errors in terms of the degree to which errors are made salient to learners. Since the methods of providing feedback included the use of colored pens, overt correction, coded correction, etc., it is reasonable to assume that the authors are referring to salience in terms of prominence or detectability rather than

in terms of importance. Nevertheless, no explicit definition of salience is provided.

Studies in perception (Odom et al., 1975; Odom and Lemond, 1975; Efron et al., 1983, and many others) consider salience to mean perceptibility. Most of those studies concern children's ability to perceive shape, color, etc. McDonald and Heilenman (1991) equate salience and detectability in their research on determinants of adult usage of syntactic and semantic cues in sentence interpretation. In addition, attention and detection literature often equates salience with detectability.[1]

Thus, when salience is defined at all, we see that the general literature in fields related to second language acquisition refers to three different notions of salience: Importance, prominence, and detectability.

Treatment of salience in the SLA literature is similar to that which occurs in the general literature. In other words, although use of the term "salience" is fairly frequent in SLA literature, this use usually occurs in the absence of any attempt at defining salience or characterizing its relationship to SLA performance.

Like researchers in the general literature, some SLA investigators define salience in terms of factors affecting it. For example, Bardovi-Harlig (1987) refers to salience as availability, i.e. frequency, of the data during learner exposure to the target language. She clearly leaves open, however, the possibility of other meanings: "salience (in

this case, the availability of data)" (p. 387). Her phrasing suggests that she is defining salience; a more logical interpretation, however, is that she proposes frequency as a factor which contributes to salience. That it is not the only factor is made clear by her use of "in this case." Furthermore, it must be noted that Bardovi-Harlig is, to the best of the present investigator's knowledge, the only researcher who cites frequency as a contributor to salience. Other investigators (see Chafe, 1977, Clark and Clark, 1977) have found that unexpectedness contributes to salience. This last observation would appear to run counter to Bardovi-Harlig's suggestion, at least if one understands unexpectedness to be the result of a *lack* of frequency. However, it is not necessarily the infrequency of an event which causes it to be unexpected; one has to consider also the context in which the event occurs. One can imagine, for example, a frequently occurring utterance such as "hello" being unexpected if uttered in response to an interlocutor's expression of thanks.[2] Since unexpectedness is not solely the result of a relative infrequency of occurrence of a structure, we see that the two claims (that frequency, on the one hand, and unexpectedness, on the other, are contributors to salience) are not necessarily in conflict.

Other SLA investigators who suggest factors contributing to salience include Strickland and Ayoun (1989), who suggest a hierarchy of importance for

communicative elements and note that content (as opposed to form, at least in their study) appears to contribute to an item's salience. They suggested that the relative lack of content in pronouns caused learners to pay less attention to sentence subjects, which were all pronouns, than was predicted. Thus they propose content as a factor influencing salience.

Galambos and Hakuta (1988) postulate that gender errors in Spanish are salient, while noun-verb number agreement errors are not salient. Their argument is based on findings from acquisition orders, where gender is acquired early and noun-verb agreement relatively late. Additionally, they comment that the high functional load carried by gender in Spanish may influence the salience of gender errors. For them, then, high functional load is a factor influencing salience. They note an effect of item salience on the metalinguistic performance of bilingual children.

Thus the language acquisition literature, like literature in other fields, skirts the issue of defining salience. When there does exist more than a mere mentioning of the term "salience" or a reference to what affects salience, discussions on the topic are sometimes of a circular nature: What is salient gets noticed, and that which is noticeable is salient. Salience is often defined in a very narrow way which differs depending on the researcher. For example, the notion "salience" might mean importance (e.g. communicative value) to one investigator, while to

another, it refers to prominence (in the sense of a subjective measure of how much an item "stands out"). Yet the same term (i.e. "salience") is employed in each case, and each researcher uses that term as if it were unidimensional, and as if there were general agreement as to its definition.

While some researchers define salience in a narrow sense, others appear to use it as if it were an omnibus or a unitary concept, i.e. a cover term for many features, all deriving from the same source. This is difficult to ascertain, however, since those features are not explicitly named.

As in the general literature, those definitions which are given cluster around three main concepts: importance, prominence and detectability. Strickland and Ayoun (1989) seem to refer to salience in terms of both communicative importance and prominence. Galambos and Hakuta (1988) appear to equate salience and prominence, although not explicitly.

Hatch (1983) defines salience as perceptual prominence. She suggests that learners incorporate free morphemes more readily than bound morphemes into their interlanguage because the former are stressed and therefore more perceptually prominent. She does not discuss salience, per se, in detail, but notes that an item may achieve added salience when a learning strategy such as Slobin's (1973)

operating principle "pay attention to ends of words" is employed.

Sharwood Smith (1991) discusses salience in terms of noticeability or "mental flags" and draws a distinction between externally induced salience and internal salience. The former is deliberately engineered by someone other than the learner. The latter, internal salience, is salience which occurs naturally "because the learner is ready for a growth in knowledge" (p.121). Sharwood Smith gives the example of a morphological inflection colored green by a teacher or textbook writer, in an attempt to create a "flag" for the learner. The hope is that learners will use such flags to develop their own internal mental flags. Nevertheless, Sharwood Smith recognizes that "what is made salient by the teacher may not be perceived as salient by the learner" (p. 120).[3]

In this section, we have seen that while discussion of factors which affect salience is plentiful, as is mention of the term "salience", relatively little attention has been given to defining salience, in either the general or the SLA literature. Why is this a problem? It is not the case that confused (and confusing) definitions of salience lead researchers to contradictory findings; as previously mentioned, salience per se has not been studied. Since it has generally not been an object of study in and of itself, the way in which salience is defined does not affect results of studies. What is problematic, however, is the use of

salience (undefined) as a means of explaining unexpected results. Both Bardovi-Harlig (1987) and Gass (1980) have suggested that the salience of a given construction (English preposition stranding and English genitive, respectively) may explain unexpectedly early mastery. Although Bardovi-Harlig defines salience for her study as availability, she leaves open the possibility that other definitions of salience might exist for her. Might, then, some other type of salience also cause unexpectedly early mastery of a construction? It is impossible to know without a more precise notion of salience. (For more discussion of Bardovi-Harlig, see Chapter 2.)

Gass does not define salience. She investigated errors and avoidance in acquisition of English relative clause types and found that genitives were mastered out of order (second only to subject relative clauses) with respect to the order predicted by the NP Accessibility Hierarchy of Keenan and Comrie (1977). Gass notes the following: 1) "whose" is the only case-marked relativizer (she evidently does not consider "whom" to be case-marked) and 2) "whose" is the only relativizer that does not allow for variation (e.g. "which", "that"). She suggests that increased salience of the genitive as a result of these facts may account for their unexpectedly high mastery ranking.

If salience is to be a useful explanation, serious attention must be given to the way in which it is defined.

In the next section, we examine additional reasons why a consideration of salience is important to SLA research

1.3. The Relevance of Salience to SLA

1.3.1. Theories of Intake

Traditionally, second language acquisition research has emphasized the linguistic environment of the learner, that is, the target language input the learner receives. Another area of focus has been on the nature of language itself. In contrast, relatively little attention has been given to the cognitive capacity of the learner (Schachter, 1986). When investigators have considered learner characteristics, discussion has typically been limited to learners' affect, motivation, personality, aptitude, etc. (e.g. Krashen, 1985, and others). Too often, SLA studies view the learner as a passive recipient of a variety of kinds of knowledge, and as someone who must be protected from a potentially overwhelming barrage of complex input. As Schachter points out, the processing-model approach, which has its roots in cognitive psychology and information-processing models, has received relatively little attention from second language acquisition researchers. (For some notable exceptions, see McLaughlin et al., 1983; Chaudron, 1985; O'Malley and Chamot, 1989.)

Recently, second language acquisition researchers have begun to address the following question: What are the cognitive prerequisites for the creation of representations

of the target language from the linguistic input that learners receive? Studies of intake (Chaudron, 1985; Liceras, 1985, etc.) provide a framework within the SLA field for a conception of the language learner as an active participant in all stages of the language learning process. Intake concerns that part of the input data which is "taken in," "up-taken," etc., by the learner. Terminology varies, but the concept is the same: Not all of the data received as input by the learner become intake. Theories of intake have their underpinnings in the notion that neither examination of the quantity of input required by learners (see Seliger, 1977; Bialystok, 1978; Swain, 1985, and others) nor of the quality of that input (see Krashen, 1983; Beebe, 1985; Lightbown and d'Anglejan, 1985, etc.) is sufficient to determine how learners acquire a second language.

Corder (1967:165) was the first investigator to note the need to draw a distinction between input and intake. He defines intake in the following way:

> input is 'what goes in', not what is available for going in, and we may reasonably suppose that it is the learner who controls this input, or more properly his intake. This may well be determined by the characteristics of his language acquisition mechanism.

Other views include an equation of intake with comprehensible input, that is, input which is understood by the learner but which contains linguistic items from the

next higher (i+1) level (see Krashen, 1985). Thus for Krashen, intake is data assimilated and used by the learner to promote linguistic development in the L2.

Sharwood Smith (1985) suggests that intake is the interpretation of outside input (e.g. a teacher's utterance) by the combined forces of the acquisition device and inside input. The latter, i.e. "inside input", appears to refer to learning strategies and awareness of the learner.

Liceras (1985) does not offer a definition of intake but proposes that it performs the role of mediator between input and output at each of many different levels of proficiency, and that cognitive factors intervene at the level of intake. The three cognitive factors she identifies are metalinguistic ability, attained linguistic knowledge, and a notion of linguistic markedness.

Chaudron (1985b:1) says that intake refers to "the mediating process between the target language available to learners as input and the learner's internalized set of L2 rules and strategies for second language development."

Most of these formulations have in common the notion that L2 learner intake is the mediator between the linguistic input received and the linguistic knowledge attained by the learner.[4] Although a starting point for a description of intake is provided by Liceras' suggestion that markedness notions, metalinguistic ability and attained linguistic knowledge affect intake, we have, thus far, no notion of what "triggers" a transformation of input to

intake. Schmidt (1990:16) points out that any adequate theory of the distinction between input and intake must consider the fact that input is external and intake is internal. He proposes the following hypothesis:

> The Intake Hypothesis: Intake is that part of the input that the learner notices.

According to Schmidt, what is noticed is taken in. Pursuing this idea to its logical conclusion, one might conclude that noticing is what triggers the transformation of input to intake. Note, however, that the term "notice" is unrevealing; it supplies no information about what processes are occurring. Since models of selective attention have long been employed in a variety of domains in cognitive psychology, it is reasonable to suggest that such models may provide a comprehensive and well-specified theoretical framework for a research domain which, according at least to Chaudron (1985b), is in need of one. Schmidt's hypothesis might therefore be revised as follows:

> Intake is that part of the input to which the learner selectively attends.

The relationship of the previous discussion to salience remains to be elucidated. Crucial to that endeavor is the notion that a selective attention mechanism may operate at several different levels. Norman (1968) suggests that there is an automatic analysis of sensory inputs and then a

selection of important inputs (via his proposed "pertinence mechanism") for further attention. His theory, then, assumes that a certain level of attention is gained by all inputs, but only those selected by the pertinence mechanism receive additional attention. Suggestions that there may exist degrees of attention have been made by other investigators (Treisman, 1969; den Heyer, 1986; Nissen and Bullemer, 1987). Schneider, Dumais and Shiffrin (1984) note that attention can be automatized. They agree that automatic control does not consume limited control-processing resources. However, they note that such a statement does not imply that automatic performance cannot benefit from allocation of control-processing resources. Posner (1980) distinguishes between two different aspects of the attention system: Orienting and detecting. The former refers to the direction at which attention is directed, while the latter refers to the contact between the input and the attentional system, such that a response can be made. A distinction such as Posner's could partially explain observations of degrees of attention.

Although she does not relate them directly either to selective attention or to salience, Liceras (1985) suggests three cognitive factors which may be operative in the intake process: 1) attained linguistic knowledge, 2) metalinguistic ability, and 3) a notion of which forms are marked and which are unmarked. These factors relate, respectively, to the novice/expert distinction, degrees of

metalinguistic ability, and the role of linguistic markedness in SLA. All of these are discussed in subsequent chapters of this dissertation.

There may be a crucial role for salience in the determination of which inputs receive further attention, and therefore an important role for salience in determining intake. One of the motivating forces of this dissertation is the notion that the salience of input may affect the operation of a selective attention device. This idea is neither new nor dissimilar to that expressed by Norman in his pertinence mechanism theory, discussed above. The crucial distinction to be drawn is that while Norman's pertinence mechanism appears to involve solely the issue of importance in determining which inputs receive further attention, the view of salience outlined in this dissertation (see section 1.4) involves not only importance, but also prominence and detectability as potential components of salience.

1.3.2. Negative Evidence

The role of negative evidence is a thorny issue for second language acquisition. In the language acquisition literature, the term "negative evidence" was, for some time, used exclusively to describe corrections of the learner's speech production (Schachter, 1991). More recently, however, studies of both L1 (cf. Demetras et al., 1986; Bohannon and Stanowicz, 1988) and L2 acquisition (Freed,

1980; Schachter, 1984) have included a wider range of phenomena under the rubric "negative evidence." Such phenomena include the following: confirmation checks (e.g. listener says "He went yesterday, right?"), expanded and corrected repetitions (i.e. recasts), clarification requests, and overt signs of incomprehension (e.g. "Huh?" or inappropriate silence).

Although the usefulness of negative evidence has been called into question for first language acquisition (v. Brown and Hanlon, 1970, and others), it has been suggested that some types of L2 learner hypotheses require negative evidence in order to be refuted (see Bley-Vroman, 1986; Schachter, 1986). Schachter discusses a situation where L2 learners hypothesized (wrongly) that free-variation existed in the target language. Second language learners of English hypothesized that infinitival "to" is optional in English. Thus they produced not only the correct sentences "I want to go" and "I can go" but also the incorrect "*I want go" and "*I can to go". Bley-Vroman also discovered free-variation hypotheses among learners, and points out that these types of hypotheses can only be disconfirmed by negative evidence.

Tomasello and Herron (1989) point out that little L2 research has been devoted to the importance of feedback. Chaudron (1977) found that, in an immediate conversational context, feedback which helped students localize errors also helped them correct the errors. Herron and Tomasello (1988) found that when class discussion of grammatical structures

centered around students' own attempts and teacher feedback (aid in localizing and identifying errors), the students learned better than when merely presented with correct examples of the structure.

Herron and Tomasello (1988) and Tomasello and Herron (1989) compared two methods for correcting errors made by students of French. In the first study, Herron and Tomasello investigated the process of overgeneralization and its correction by teaching, in one of two different ways, various "exceptions to the rule" in French grammar. One of the teaching methods involved a traditional presentation of the rule and its exceptions. The other method allowed students to induce the rule from examples given and then to make overgeneralization mistakes on the exceptions. This second method was called the "Garden Path" technique. The investigators found that the latter technique produced superior student learning and that this advantage lasted. They suggest that the Garden Path method of correction may have helped both in drawing students' attention to the rule itself--by requiring them to generalize--and in drawing their attention to the exception.

Except that they involve correction, all of these methods of feedback detailed above are analogous to Sharwood Smith's (1991) "input salience creation", in which teachers raise "flags" in an attempt to cause learners to raise their own mental flags. How successful these attempts are may in part depend on learners themselves. The fact that

individual differences exist in learners' use of negative evidence is discussed by Birdsong (1989). Some individuals have been found to use negative evidence successfully in order to modify or discard a hypothesis, while other individuals persist in a behavior in spite of receiving negative evidence regarding that behavior (see Wason, 1968; Giacobbe and Cammarota, 1986, cited in Birdsong, 1989). Thus, the salience of a particular construction involved in a piece of negative evidence is obviously not the only factor which may determine the success of the evidence. Nevertheless, elucidation of which types of constructions, and which types of errors in those constructions, are more salient to learners will enable instructors to be more effective in their provision of negative evidence.

1.3.3. Error Evaluation

Language learners make errors. Proponents of cognitive approaches to second language learning (e.g. Corder, 1967; Selinker; McLaughlin, 1987; Bley-Vroman, 1986) have emphasized that errors are a natural occurrence in developing proficiency. In their view, errors provide clues about the hypotheses currently being entertained by second language learners.

Evaluations of the errors learners make can be equally informative. Second language acquisition (SLA) research on error evaluations has dealt with a variety of issues: 1) native speaker reaction to errors (Politzer, 1978; Chastain,

1980; Piazza, 1980; Delisle, 1982; Ensz, 1982, etc.), 2) teachers' reactions to learner errors (Birdsong, Thoren, and Price, 1984; Birdsong and Kassen, 1988, etc.), 3) uniformity of error rating between groups of informants (Ervin, 1978; Galloway, 1980), and 4) the best methods of error correction (Cathcart and Olsen, 1976; Magnan, 1982). Of all of these studies, only two (Cathcart and Olsen, Birdsong and Kassen) involved students as informants. According to Birdsong and Kassen, "a significant group of error evaluators has been systematically overlooked: students." (p.1) They point out that learners' error evaluations are particularly significant when compared to those of their instructors, since agreement as to the grievousness of a given error might lead to more effective remediation of it.

Schmidt's suggestion that input must be noticed in order to become intake was discussed above. (See section 1.3.1.) Even if one rejects Schmidt's (1990) argument that second language learning does not take place without noticing, it is reasonable to suggest that in order to be evaluated, an error must first be noticed. Therefore, at least as crucial as student error evaluation data is information about what types of errors learners notice. The former is underrepresented in SLA literature. To the best of my knowledge, data on learner patterns of noticing errors are completely absent. Data on the salience of errors to learners are equally nonexistent, although Birdsong and

Kassen (1988) refer to a possible effect of salience on (within group) learner judgments of severity. One of the aims of this study is a partial filling of this gap in current knowledge of learner error evaluations in general and of the salience of errors to learners, in particular. The specific research questions addressed in this dissertation are outlined in the next section.

1.4. Statement of the Problem

In the last section, we saw that SLA literature currently contains few data regarding learner error evaluations, no data on learner error detection, and few reliable data about what is salient to learners of a second language. The last observation, that of a lack of data regarding what is salient to learners, is not surprising, given the current lack of consensus (and, indeed, of discussion) about what it means to say that an item is "salient."

In this chapter, we have seen that definitions of salience, where they exist at all, appear to consistently involve three concepts: importance, prominence, and detectability. It is unclear whether salience is best viewed as a cluster of distinct components or as a unitary concept consisting of interrelated components. This dissertation proposes that importance, prominence, and detectability are all possible components of salience, and seeks to discover whether they are distinct (i.e.

independent) components, or whether they are instead components which covary. Thus one of the issues explored in the current study is that of whether salience is a unitary concept.

In addition to the impreciseness and variability exhibited in definitions of salience, few researchers entertain the notion that the salience of an item to a learner may change with the learner's experience (see Chaudron, 1983; Birdsong, 1989; Sharwood Smith, 1991, as exceptions). In other words, many investigators appear to assume that salience is a stable concept. Since this assumption may not in fact be correct, there exists a potential for addition to the confusion which is already created by the lack of a precise definition of salience. Accordingly, this dissertation explores the question of whether salience is a stable concept.

Meaningful discussion of salience with regard to second language acquisition encounters a third obstacle: Virtually nonexistent are claims as to what constructions (or errors within those constructions, depending on the focus of the study) are the more salient ones for second language learners. By making use of a few notable exceptions to this trend (these exceptions are discussed in Chapter 2), the current study explores a third issue; namely, whether some types of errors are more salient than others. Salience of error types is investigated along each of the proposed

components of salience: importance, prominence, and detectability.

This experimental study examines the salience of French errors to three groups of native English speaker subjects: beginning students, intermediate students, and nonnative teachers of French. The major questions to be addressed in this dissertation are the following:

1) Do the three types of proposed salience components (i.e. importance, prominence, and detectability) correlate with each other, overall? Do they correlate when subdivided by subject group and error type? Is response to a given type of error consistent across the three salience components? The issue: Is salience a unitary concept, i.e. can it be meaningfully discussed without reference to several distinct components?

2) Is there an effect for level of experience with French, such that error types differ in their detectability, importance and prominence, as a function of the learner's experience and/or expertise? The issue: Is salience a stable concept?

3) Is there a main effect for the type of linguistic construction in which the error occurs? The issue: Is inherent salience of a given type of error a reasonable notion for second language acquisition?

This chapter addressed the relevance of salience to second language acquisition and reviewed existing notions of salience and problems involved in characterizing it for second language acquisition. We now turn to Chapter 2, in which each of the research questions posed in this dissertation is discussed.

Notes

1. Although some researchers might argue that these last two notions, perceptibility and detectability, are identical, a distinction between them seems to have been based primarily on the identity of the object of investigation. For example, one "perceives" shape, yet "detects" errors. Such a distinction is not crucial to the current study; the term "detectability" will be employed here.

2. In addition, it has been noted (Oller, 1988) that inappropriateness of an utterance may have an undesirable result (e.g. embarrassment) which may in turn lead to eradication of the error in question.

3. The question arises here as to whether discussion of the salience of an item has any meaning without reference to a person to whom the item is more or less salient. In other words, "made salient by the teacher" invites the question "To whom?" This issue is discussed in Chapter 2.

4. The "attained linguistic knowledge" to which Liceras refers (i.e. one of the cognitive factors which influences intake) is that knowledge already possessed by the learner, while the attained linguistic knowledge mentioned here is that gained as a result of intake. It may appear circular to claim, in effect, that attained linguistic knowledge affects intake and results in attained linguistic knowledge. However, it must be remembered that, in most cases of L2 acquisition, attained linguistic knowledge at time x is not the same as it is at time y; the nature of such knowledge is that it is not static.

CHAPTER 2
EXAMINATION OF THE RESEARCH QUESTIONS

2.1. Is Salience a Unitary Concept?

2.1.1. General Remarks

In Chapter 1, we saw that salience is not often defined in the literature, but that when it is, such definitions usually mention detectability, prominence, or importance. Accordingly, the present dissertation proposes these three concepts as possible components of salience. As mentioned in Chapter 1, the question then arises as to whether these components are distinct or whether they are interrelated. In other terms, do detectability, prominence and importance covary? An answer to this question should provide evidence for deciding whether salience is best viewed as a cluster of distinct components or as a unitary concept consisting of interrelated components.

Before discussing predictions regarding this question, it is necessary to examine in broader terms each of these proposed components to see why it makes sense to suggest that they may be components of salience.

2.1.2. Detectability

In the present study, the detectability of a given error to a subject is operationally defined as the speed with which the error is found by that subject.[1]

Detectability is thus measured by subject reaction times on the error detection task (see Chapter 3).

A task such as the error detection task, in which a subject decides whether or not a sentence contains an error, requires the separation of an existing error from the material in the rest of the sentence. In this way, the error detection task is like any other task which requires the separation of figure from ground (see section 3.1.5). The more readily (i.e. quickly) the error is separated out from the unflawed portion of the sentence, the more detectable it is said to be.[2]

Relatedly, it is possible to view detectability from the broader perspective of signal detection theory (see Green and Swets, 1964). The signal detection task differs from a classical psychophysical experiment in two ways: 1) only one stimulus intensity is presented, and 2) on some of the trials, no stimulus is presented. Subjects are asked to say whether they heard a tone ("yes") or not ("no"). In terms of signal detection theory, however, the task is actually to decide whether a presentation contains a signal plus noise (S+N) or merely noise (N). Subjects' correct responses of "yes" and "no" (known as "hits" and "correct rejections," respectively) are tabulated, as are incorrect responses of "yes" and "no" ("false alarms" and "misses"). This treatment of the psychophysical problem of stimulus sensitivity stems from a conviction that the notion of an absolute threshold of sensitivity (the classical idea) is

not a valid one, since factors other than the subject's sensitivity to the signal may influence the results of psychophysical experiments. One crucial factor is the subject's response criterion, that is, the subject's criterion for indicating perception of the signal. This response criterion varies between subjects and is also vulnerable to within-subject variation. Signal detection theory is designed to handle such variability through the use of probability statistics.

Of course, placement of the error detection task within the broader context of the signal detection task is possible only in a limited sense; the present study is not a psychophysical one, and response criteria are not manipulated. Nevertheless, the subject who performs an error detection task is faced with deciding whether a given sentence contains a signal (in this case, an error) in the midst of noise (the rest of the sentence) or whether the sentence contains only noise (i.e. it is a correct sentence). Birdsong (in progress: 2) notes that signal detection theory concerns itself with the parameters of judgment and decision-making under conditions of uncertainty. He then points out that since, according to this theory, judgments are viewed as the product of both sensory and cognitive mechanisms, it follows that signal detection theory "allows for generalization to a number of performance domains."

The use of detectability to define salience is widespread in the literature, as has been pointed out. The fact that detection does not occur in a vacuum but rather in the midst of "noise" is sometimes overlooked. Similarly overlooked is the fact that discussion of the salience of a given item (to a perceiver) only has meaning when the presence of other items is assumed. In other words, the phenomenon of salience does not occur in a vacuum, either. It is in their shared notion of "separating out" that detectability and salience can be seen to be related.

2.1.3. Prominence

Closely related to detectability is the notion of prominence. The prominence of an error to a subject is operationally defined in this study as the degree to which the subject finds the error to be "glaring." Prominence is measured on a rating scale from 1 ("not at all glaring") to 6 ("very glaring").

Like measures of detectability, prominence measures concern the subject's ability to separate out the error from the rest of the sentence. Where the two measures differ is in degree of subjectivity; prominence ratings reflect (presumably accurately but not, it is true, necessarily so) the subject's *perceived* ease in detecting a given error. Whether the two measures correlate will be of much interest (for more discussion, see section 2.1.5).

2.1.4. Importance

Importance measures, like prominence measures, are subjective in nature. The importance of an error to a subject is operationally defined as the degree to which the subject finds the error to be important to correct (see section 3.1.4). Importance is measured on a rating scale from 1 ("not at all important") to 6 ("very important").

In addition to the appearance of importance in the literature as a synonym for salience, use of the importance measure, particularly as it is defined in this study, is motivated by Birdsong and Kassen's (1988) finding that learners and teachers agree on the importance[3] of a given error relative to that of another error. Of interest, among other questions, will be whether these results are replicated and whether importance measures correlate with those of detectability and prominence.

2.1.5. The Research Hypotheses

The research question we have been addressing is the following: Is salience a unitary concept? In order to answer it, we must first discover whether the proposed salience components of detectability, prominence, and importance correlate.

The first hypothesis related to this question emerges partly from the discussion above, in which was emphasized the closeness of the relationship between detectability and

prominence, and partly from speculation that subjects' perceptions about how detectable a given error is to them are reasonably accurate. Hypothesis 1 predicts that measures of detectability and prominence will correlate positively.

There is no well-principled rationale for predicting a correlation either between prominence and importance or between importance and detectability. In fact, it is easy to imagine that, say, omission of definite articles might be an error which is both prominent and detectable to native English speakers, but which is judged by them to be unimportant, especially if "importance" is interpreted to mean communicative importance. Hypothesis 2 predicts a lack of correlation between measures of prominence and importance, and Hypothesis 3 predicts no correlation between importance and detectability measures.

Note that if all of these hypotheses are supported in results of the study, we will conclude, at least in broadest terms, that salience is not a unitary concept. If Hypothesis 1 is confirmed, however, then that result will suggest unitariness, at least along the dimensions of detectability and prominence. We turn now to a consideration of the second research question.

2.2. Is Salience a Stable Concept?

2.2.1. General remarks

A second question addressed in this research is that of whether salience is a stable concept. In the previous chapter, we stated that few researchers consider the idea that the salience of a linguistic item to a learner may change along with, and as a result of, the learner's experience with the target language. While it is reasonable to expect amount of exposure to the L2 to affect the salience of errors to subjects, it is also reasonable to assume that individual differences between subjects, both in expertise (see distinction drawn below between exposure and expertise) and in metalinguistic performance level, may have an effect. Both of these issues are discussed in this section.

2.2.2. The Novice/Expert Distinction

The distinction between novices and experts applies to many domains of skilled performance. Studies have shown that experts use different strategies and techniques for representing information than do novices (Nation and McLaughlin, 1986). For example, Adelson (1981, 1984) has reported that expert programmers use abstract representations when trying to recall programming material, whereas novices use more concrete representations. Chase and Simon (1973) found that, after only a 5 second viewing of midgame boards, master chess players can reconstruct

those boards with higher than 90% accuracy. Chase and Simon observed that expert players recalled configurational clusters, whereas beginners lacked the skill to form these types of abstract representations. Nation and McLaughlin (1986) suggest that a similar distinction may be drawn between experts and novices in the domain of language learning; that is, that experts may form abstract conceptualizations which novices are not capable of forming. More will be said about this shortly.

One element crucial to the novice/expert issue is a distinction drawn by information processing theories between automatic and controlled processing. The theory (Shiffrin and Schneider, 1977; Schneider, Dumais and Shiffrin, 1984, and others) is that some processes become routinized through practice, to the point where very little cognitive effort is required in order to perform them. These are called automatic processes, and performance of them is assumed to require little attention and to take place without consuming (limited capacity) resources. From this perspective, experts can be viewed as having achieved routinization of tasks which, for novices, require attention and effort to perform.

As was briefly hinted at above, application of information processing theories to second language acquisition (Nation and McLaughlin, 1986; Bialystok and Ryan, 1985, among others) has resulted in the notion that language learners, also, differ in the ways in which they

process information, depending on their level of expertise with the target language, or with languages in general. Nation and McLaughlin (1986) propose that "'expert'[4] language learners do not use the same information-processing strategies and techniques as do more 'novice' learners" (p.42). They define "experts" as multilinguals and "novices" as monolinguals or bilinguals. In their study, they find that experts learn an artificial grammar significantly better than novices under implicit learning conditions (i.e. where subjects are not asked to form rules), but that the experts fare no better than the novices under explicit learning conditions (where they are instructed to discover rules). This finding leads the investigators to hypothesize that the superior performance of the experts may derive at least partially from a tendency to create abstract representations of the material even when given no instructions to derive underlying rules for that material.

The idea that experience and expertise affect what is salient to a learner has been mentioned in literature in the domains of child perception (Odom and Lemond, 1975), first language acquisition (Peters, 1985) and second language acquisition (Chaudron, 1982; Birdsong, 1989). In a study of the effects of variations in topic reinstatement on ESL learners' comprehension of oral text, Chaudron (1982) tested learner recognition and recall of five different linguistic devices: 1) Simple noun (simple topic reiteration, 2)

synonym, 3) repeated noun, 4) topicalizing rhetorical question, and 5) If-Clause (nonconditional). The structures chosen were considered to vary not only in linguistic complexity (increasing from 1 to 5), but also in their psychological salience as reinstatement devices. Combining both the simplicity and the salience factors, the following ordering of effectiveness of devices was predicted:

Rhetorical Question > Repeated Noun > If-Clause > Simple Noun > Synonym

The most important finding with respect to the present study is that while the low-skill learners (as determined by scores on a standardized test) performed less well on recall of each of the devices than did the middle- or high-skill groups (i.e. the devices were less salient to them), the low-skill group had relatively less difficulty with the synonym than the other two groups did. Chaudron suggests that the synonym was relatively less salient to the higher skill listeners because they were more engaged in processing the synonym into a mental representation, while the low-skill listeners were recalling surface features of the text. This suggestion is consonant with Rossman's (1981) finding that native speakers attended more to meaning than to form in a reading task, while the opposite obtained for nonnative speakers.

Since the present study does not test recall, we can expect not to find a situation such as that outlined above with respect to the synonym device in Chaudron's study. Nevertheless, it is reasonable to expect expertise to have a positive effect on error detection performance. Birdsong (1989) makes this point. He cautions, however, that error detection ability is subject to individual differences, and should not be considered uniquely a function of either exposure or expertise.

It is important to draw a clear distinction between exposure to the target language and expertise with that language, because the two do not necessarily correlate. Acquired expertise with a target language is a function of much more than merely exposure; this fact has been demonstrated repeatedly in SLA research (for more discussion, see endnote 3). Any study which uses L2 exposure as an independent variable must pay careful attention to the problem of what constitutes a novice learner and what an expert one. More will be said about this in subsequent chapters. It should be pointed out that although discussion in this section (and indeed, in most of the literature) refers to novices and experts as if there were a clearly drawn line between them, i.e. as if a clear dichotomy were involved, such a representation is adopted more for simplicity of exposition than for its basis in reality. Discussion of the novice/expert distinction in terms of degrees of expertise, represented on a continuum,

is clearly the more appropriate, if not always the more convenient, means of examining the issue.

A suggestion made in this dissertation is that the relative salience value of an error for a second language learner may change as that learner gains expertise with the language in question. So far, we have seen some evidence that differences may exist (both between groups and within groups) in error detection ability. The expected similarity between error detection and ratings of prominence has been detailed above; what, however, can we say about the likely effect of expertise on subject ratings of importance? We do, after all, require evidence from each of the tasks performed in order to answer the research questions.

In a study of uniformity of error severity judgments between teachers and students, Birdsong and Kassen (1988) provide a rationale for hypothesizing that congruence in judgments may be found. They conjecture that, in cases where English natives are teaching French to other English natives, a shared linguistic background and similar (classroom) learning experiences might cause teachers and students to share error judgment rationales. Another factor they suggest may influence degree of judgment similarity is the length of target language exposure experienced by students. The reasoning behind this is that teachers, either explicitly (e.g. correction) or implicitly (e.g. facial expression, etc.) react to errors and, over time, students' reactions may standardize.

Errors in the Birdsong and Kassen study were of four types: morphological, syntactic, lexical and phonological. All were errors of the sort frequently made by beginning students and frequently corrected in class; recognition of these errors did not require intuitive knowledge of the language.

Results from the Birdsong and Kassen study indicated, first, the importance of distinguishing between severity of judgment and relative seriousness of error (see Chapter 3 for a discussion). The result pertinent to present discussion is that while there was agreement among teachers and both groups of students as to how important a given error was relative to another, the judgments of more experienced students (i.e. those with greater length of exposure) matched those of the English native speaker teachers more closely, in magnitude, than did the beginner judgments.

In this section, we have seen evidence that the novice/expert distinction, in the sense of both degree of expertise and length of exposure, may have an effect on subject performance. We have also seen that attachment of the labels "novice" and "expert" to subjects carries with it a danger not only of attempting to create a dichotomy where it almost surely does not exist, but also of ignoring the role of individual differences in second language learning. Such differences may underlie the novice/expert distinction. Many other kinds of individual differences exist and have

been investigated (Nation and McLaughlin, 1986; Birdsong, 1989, and numerous others). One of these differences is found in learners' ability to perform metalinguistic tasks. We will first define metalinguistic performance and then discuss individual differences in such performance.

2.2.3. Metalinguistic Performance

The term "metalinguistic" is used in the literature in combination with a variety of other terms: Ability, awareness, knowledge, skill, performance, etc. (see Birdsong 1989 for a thorough discussion). In the interests of maximum concreteness, discussion here centers on the last, i.e. performance, since it lends itself to direct observation in a way that the others do not.

One frequent definition of metalinguistic performance is "attention to linguistic form rather than content" (Birdsong, 1989:1). Cummins (1978) says it involves looking "at language rather than through it to the intended meaning" (p.127). Ryan and Ledger (1979) explain metalinguistic performance as a shift of attention from meaning to structure during tasks which require deliberate control over language forms.

Tasks which have been deemed metalinguistic include (but are not restricted to) the following: Recognition of ambiguity, at both the syntactic and the lexical level (e.g. the comprehension of ambiguous sentences and of puns), recognition of sound patterns, such as those involved in

rhyming games, poetry, and the noticing of regional or foreign accents, identification of structural deviance (e.g. proof-reading, making grammaticality judgments), adaptation of speech to a listener's needs, use of figurative language, etc (see Birdsong, 1989).

Bialystok and Ryan (1985) point out that there is no current consensus on either the precise domain of activities which may be termed metalinguistic or the age at which metalinguistic performance may first be said to be taking place. Similarly, Kamhi (1987) states that there is a lack of agreement among investigators regarding whether or not certain of the tasks mentioned above are, in fact, properly considered to be metalinguistic in nature. Nesdale and Tunmer (1984) emphasize a present lack of refined and generally accepted investigative methods in the domain of metalinguistic research. "In most areas of research on children's metalinguistic awareness at the present time, it is unclear whether conflicting estimates revealed by different studies are due to sampling, task, procedural or measurement differences." (p. 36). The authors point out that some of these difficulties are encountered by any researchers working with children, but that others are unique to metalinguistic research. They do not elaborate further on the nature of these unique difficulties, but it seems intuitive that the lack of a refined, generally accepted idea of what one is investigating is reflected in a lack of refined, acceptable methods for investigating it.

Even where there is agreement that a given set of tasks are all metalinguistic in nature, a high level of performance on one of them does not necessarily correspond to a similar level of performance on another. As Birdsong (1989:48) points out, this makes intuitive sense: "...one can imagine people who are gifted punsters yet poor poets". Scribner and Cole (1981) caution against viewing metalinguistic performance in a unitary fashion.

Bialystok and Ryan (1985b) note that past definitions of metalinguistic performance have emphasized either structured knowledge of the language (e.g. awareness of semantic or syntactic language properties, knowledge of functions and features of language, etc.), control over language form (e.g. ability to attend to language forms, ability to shift attention from meaning to form, etc.) or a combination of structured knowledge and control (e.g. thinking about language, manipulating language as an object, etc.)

Bialystok and Ryan argue that, although these two aspects of metalinguistic performance are interrelated, distinctions need to be made between them not only in order to clarify the nature of metalinguistic development, but also to provide for a better understanding of a wide range of linguistic and metalinguistic developments which occur both in first and second language acquisition. Accordingly, they propose a framework which assumes that a minimal but common set of underlying cognitive skills is

responsible for language proficiency. They treat analyzed knowledge and control as linear dimensions which characterize both tasks and abilities in terms of degrees (i.e. avoiding dichotomies).

According to their metacognitive framework, then, metalinguistic performance involves the deployment of two basic cognitive skills: analysis of knowledge of language, and control over cognitive functions. The former results from the categorization of linguistic knowledge and the latter from the allocation of attentional resources. Originally (see Bialystok, 1978), the dimension of analyzed knowledge involved a distinction between knowledge that is intuitive, or implicit, and knowledge that is explicit (see Reber, 1976; Reber and Lewis, 1977). Such a distinction can also be made in terms of unconscious vs. conscious knowledge (see section 1.3.2 for a brief discussion of the role of consciousness in L2 learning), although Bialystok and Ryan (1985b) do not view consciousness as a criterion for determining analyzed vs. unanalyzed knowledge, but rather as an outcome of more advanced levels of analyzed knowledge. Furthermore, in their more recent formulation, Bialystok and Ryan view the analyzed/unanalyzed distinction as a continuum rather than a dichotomy. They suggest that metalinguistic performance entails a cooperation between analyzed knowledge and cognitive control. (For an elaboration of their framework, see Chapter 4.) In their view, literacy is the primary factor contributing to the former, while

bilingualism and schooling are the major determinants of the latter.

An additional factor of individual differences also plays a role in metalinguistic performance, however; we now examine that role.

2.2.3.1. Individual Differences in Metalinguistic Performance

The observation that some individuals perform metalinguistic activities better than others initially interested primarily researchers in first language acquisition (e.g. Ben Zeev, 1977; Scribner and Cole, 1981; Tunmer and Herriman, 1984; Bialystok and Ryan, 1985; Menyuk, 1985, etc.) The greatest concentration of studies has investigated the relationship between aspects of children's linguistic development (e.g. literacy, schooling, bilingualism) and their metalinguistic development. With regard to these aspects, arguably the most researched domain has been that of the (positive) correlation between bilingualism and enhanced metalinguistic development. Ben Zeev (1977) suggests that mutual interference between two languages forces bilingual children to develop what she terms "coping strategies" which accelerate cognitive growth. In broad terms, these strategies involve 1) an increased tendency to look for structure and 2) a willingness to reorganize perceptions. In her tests, bilingual children approached syntactic tasks in a more analytic manner than

did monolinguals. They looked for evidence of systematicity (i.e. underlying dimensions in patterns) and made more hypotheses. She concludes that bilingualism engenders greater cognitive flexibility. Ianco-Worrall (1972) attributes the superior metalinguistic performance of bilingual children primarily to an early discovery of the arbitrary nature of language. Galambos and Hakuta (1988) discovered that native language proficiency as well as a child's degree of bilingualism had an effect on metalinguistic performance.

Studies of adult metalinguistic performance have concentrated on relating such performance to L2 acquisition (Bialystok, 1978; Chaudron, 1983; Birdsong, 1989, and others). Research in adult metalinguistic performance has involved tasks similar to those employed in investigating children's metalinguistic performance, with the important difference that the focus has typically been on subjects' grammaticality judgments (Birdsong, 1989; Chaudron, 1983). The use of native speaker intuitions about the grammaticality of sentences has a tradition in linguistics dating from Chomsky's (1965) idea that such intuitions are reflections, albeit indirect, of a speaker's competence. Although grammaticality judgments have been those most extensively elicited, other tasks have included judgments of acceptability, ordinariness, ambiguity, meaningfulness, comprehensibility, synonymy, etc. (Chaudron, 1983.)

Although investigation of adult metalinguistic performance does not involve the types of age-related difficulties encountered in research with children (e.g. limited linguistic, memorial and cognitive abilities, especially in very young children; see Nesdale and Tunmer, 1984), there remain nevertheless many factors which could potentially affect adult performance. Chaudron lists variables such as age, cognitive development, training, language background, social experiences, and the context of presentation of stimuli.

In addition to *inter*subject variability which cannot be accounted for in metalinguistic terms, per se, there also exists considerable *intra*subject variation attributable to effects of repetition, order and mode of presentation, context, the presence or absence of a mirror facing the subject, etc. Any of these effects may, for example, cause a subject to change a grammaticality judgment. (For a comprehensive review, see Birdsong 1989.)

Judgments of grammaticality appear to be particularly unstable. Nagata (1988) discovered that repetition treatment makes subjects' judgment criteria more stringent for both grammatical and ungrammatical sentences. Thus the same subject may give a different judgment at Time 2 than (s)he gave to the same sentence at Time 1. Such evidence suggests that practice may be a factor.

Even if performance on grammaticality judgment tasks is no more variable than that on any other task deemed

metalinguistic (an arguable notion), excessive reliance on such performance data is methodologically unwise. Chaudron (1983) warns against taking such data too seriously. There are, in fact, other arguments to be made against a concentration on grammaticality judgments as the "prototypical" metalinguistic task. The most obvious one is the constraining influence of such a concentration. Given the wide range of operations considered to be metalinguistic in nature, an investigation of only one of them seems unnecessarily restrictive. Furthermore, present lack of agreement on what exactly constitutes metalinguistic performance makes rigorous investigative procedures mandatory. The use of four different tasks in the current research is regarded by this investigator as one of its strengths. As mentioned above, converging evidence from multiple sources provides the safest and most convincing data. Such evidence is particularly crucial in the area of metalinguistic research, where performance data on any given task are so notoriously unreliable.

Chaudron (1983) argues that it is important to distinguish metalinguistic judgments, which are a response to a stimulus for which there is (at least, in principle) a correct response, from what he terms metalinguistic sentiments (attitudes, opinions, preferences, etc.), for which there is no one correct response. He confines his review to studies of metalinguistic judgments. The distinction he discusses, between judgments and sentiments,

is analogous to a distinction between objective and subjective judgments, respectively, which has been made earlier in this chapter and will be mentioned again in Chapter 3. Although terminology differs, the spirit of the distinction is the same. This distinction is relevant to the present study, since it can be seen that, of the four tasks performed by subjects, two of them (error detection and sentence matching) are objective tasks, and the other two (judgments of error prominence and importance) are subjective tasks. All, it must be noted, are forms of metalinguistic performance and, as such, subject to individual differences. These individual differences, as well as differences in subject expertise with the L2, may potentially complicate interpretation of the research results.

2.2.4. The Research Hypotheses

The second research question is the following: Is salience a stable concept? To answer it, we must discover whether significant differences occur between the three groups (beginner, intermediate, and teacher) with respect to error detection, prominence and importance measures. From the preceding discussion emerge three hypotheses. The first is that the groups will differ in error detection ability, with teachers being the fastest and beginners the slowest. Correspondingly, we predict that differences will also obtain for prominence measures, such that teachers, relative

to students, will find errors to be more prominent. For both measures, intermediates are predicted to fall in between. The last hypothesis is that groups will not differ in relative ratings of error importance. The suggestion discussed earlier, that students derive their notions of error importance from teachers, seems logical; we expect, therefore, to replicate the findings of Birdsong and Kassen.

Hypotheses 1 and 2, then, predict that salience is not a stable concept. Hypothesis 3 predicts the opposite.

2.3. Is Salience of Types of Errors a Meaningful Notion for SLA?

2.3.1. General Remarks

The third question posed in the present research is that of whether the notion of inherent salience, irrespective of expertise or amount of L2 exposure, is a meaningful one. In other words, are some types of errors more salient than others to all subjects? Note that one interpretation of highly stable inter-group performance on the dependent measures (i.e. lack of significant between-group differences with respect to reaction time and error judgments) would be inherent salience. Therefore, the answer to the third research question may be seen to depend partly upon the answer to the second question, i.e. the stability issue.

As previously mentioned, present focus on errors is motivated by the fact that most learners of a second language are surrounded by error-ridden input (cf Birdsong

and Kassen,1988, for a discussion). Focus will center on errors involving constructions about which claims of salience have been made either in the SLA literature or in this dissertation (based on evidence from the literature). These constructions include preposition stranding, the double object construction, and relativized resumptive constructions.

2.3.2. Discussion of the Relevant Linguistic Constructions

Bardovi-Harlig's (1987) study of learners of English as a second language investigated the acquisition of preposition stranding, a putatively marked construction, and preposition pied piping, its unmarked counterpart, in both questions and relative clauses (see examples below). Following Van Riemsdijk (1978), Bardovi-Harlig identified preposition stranding as marked and preposition pied piping as unmarked, based on two arguments: 1) Typologically (i.e. cross-linguistically), preposition stranding is rare; only Indo-European languages (and not even all of these) allow it. 2) Syntactic (government and binding) theory argues that extraction of an element of a prepositional phrase (PP), which is what occurs in preposition stranding, is usually impossible, while movement of an entire constituent, as effected in preposition pied piping, is unexceptional.

Like many other SLA investigators (e.g. Mazurkewich, 1984, 1985; White, 1987; Hawkins, 1987), Bardovi-Harlig discusses markedness in terms of Chomsky's Universal Grammar

(UG). According to UG theory, rules of the core grammar are considered to be unmarked, while those of the peripheral grammar are marked. The prediction is that core (unmarked) rules can be learned on the basis of minimal language exposure, but that peripheral (marked) rules must be learned via positive evidence of their existence in the grammar. In the absence of evidence needed to choose marked options, unmarked ones are selected. Following this theory, then, Bardovi-Harlig notes that preposition pied piping ought to be acquired before preposition stranding. However, the opposite acquisition order occurred. In other words, 1), below, was acquired before 2).

(1) Preposition Stranding (marked)

Q: Who did John give the book to?

RC: The man who Cathy gave the book to was Kevin.

(2) Pied Piping (unmarked)

Q: To whom did John give the book?

RC: The man to whom Cathy gave the book was Kevin.

(Examples taken from Bardovi-Harlig, 1987.)

This result provided an apparent counterexample to the markedness hypothesis, which states that unmarked forms will be acquired before marked ones. She proposed that salience,

defined by her as availability (i.e. frequency) of data, was a factor in determining the acquisition order.

In a study of English native speakers learning French as a second language, White (1987) elicited grammaticality judgments of two marked structures, preposition stranding and double object dative constructions. In English, a dative noun phrase (NP) may appear either in a prepositional phrase, as in (3) or as the first NP of a double object construction, as in (4).

(3) Dative [NP PP]
William sent a memo to Isabel.

(4) Dative [NP NP] (double object construction)
William sent Isabel a memo.
(Examples taken from Mazurkewich, 1984.)

Both preposition stranding (see (1), above) and the double object construction (see (4), above) are grammatical in English but ungrammatical in French. In White's study, learners exhibited a much higher rejection rate for preposition stranding than for the double object construction. White does not mention salience in her discussion of the results. Consider, however, the following facts. Preposition stranding is virtually always chosen over pied piping in spoken English (indeed, use of the latter makes the speaker sound "stilted"). The same can not

be said for oral use of the double object construction [NP NP] compared to the dative [NP PP] construction. Imagine an English native speaker learning that preposition standing is not permitted in French. Since preposition stranding is the only form (s)he uses in English, at least orally, there is a large difference between the two languages. In contrast, the same learner uses the double object construction, which is not permitted in French, only part of the time in English. The other part of the time, the learner uses the dative NP PP in English (e.g. I gave the phone to my sister), which is also correct in French. Thus the two languages diverge less clearly with respect to treatment of the dative.

The point of the above discussion is that for English native speakers learning French, it may be that the absence of preposition stranding in the target language is highly detectable (or, in a subjective task, highly prominent) as a result of the pervasiveness of preposition stranding in English, the native language. To suggest that an absence could be detectable or prominent might, on the face of it, seem strange. However, it is logical that an item which is normally omnipresent should be missed when it is suddenly not there.

What is suggested above, then, is that the frequency of preposition stranding relative to pied piping in the native language (English) may increase the detectability and prominence of preposition stranding errors in the target

language (French). Kellerman (1983) discusses the learner's "psychotypology", which is the learner's perception of the relatedness of the L2 to the L1 in a given structure domain. It may be that learners perceive French to be quite different from English with regard to preposition stranding.

As mentioned above, both preposition stranding and the double object construction are grammatical in English but ungrammatical in French. In contrast, relativized resumptive pronoun constructions such as those found in the present study (and in 1 and 2, below) are ungrammatical in English as well as in French. Birdsong et al., 1984, investigated second language learners' grammaticality judgments for grammatical relativized sentences and four types of ungrammatical relativized sentences in French:

1) Resumptive Pronoun

*Il a casse les lunettes dont il en a besoin.

*He broke the glasses which he needs them.

2) Resumptive Pronoun + Preposition Stranding

*Il a casse les lunettes qu'il lit avec elles.

*He broke the glasses that he reads with them.

3) Preposition Stranding

*Il a casse les lunettes qu'il lit avec.

He broke the glasses that he reads with.

4) Preposition Stranding + Deletion of Relativizer

*Il a casse les lunettes il lit avec.

He broke the glasses he reads with.

These relativized sentences have been listed here in order of highest acceptance rate by subjects; i.e. Resumptive Pronoun sentence types were most often judged to be grammatical. It may be that resumptive pronoun errors are not as detectable or as prominent (which term is applicable will depend on whether the task is an objective or a subjective one) as stranding errors. After all, resumptive pronoun errors typically involve one extraneous word, a pronoun. Pronouns have been identified as words which are low in content; correspondingly, it has been suggested that low content words have low imagery and lack salience relative to other, higher imagery words (see Strickland and Ayoun, 1989; VanPatten, 1984; Paivio et al., 1968). It is true that prepositions are also low content words. However, a preposition stranding error usually involves more than just one word; an entire phrase is affected. The issue of how global an error is may reasonably be expected to affect its detectability or prominence.

What effect might we expect given that the relativized resumptive pronoun constructions in the study are ungrammatical in English as well as in French? Recall the Birdsong and Kassen (1988) hypothesis that shared linguistic background among subjects might produce similar importance ratings. One of the reasons for this is that subjects might relate errors in the target language to comparable errors in

their native language. If such comparison occurs, we will expect subjects to find relativized resumptive clauses (ungrammatical in English) to be more important than either preposition stranding or double object dative constructions (both of which are grammatical in English).

Several hypotheses result from the discussion presented above. The first is that preposition stranding errors in French will be most readily detectable and judged most prominent, with the double object construction and relativized resumptive clauses following, in that order. The second hypothesis is that relativized resumptive clauses will be judged most important. It seems likely that preposition stranding errors will be judged next most important, and the double object construction least important. This last prediction is a result of the frequency distributions (discussed earlier in this chapter) for preposition stranding and the double object construction in English.

The hypotheses have now all been made; for clarity, we review them in the next section.

2.4. Review of the Research Hypotheses

2.4.1. Hypotheses--Unitariness

Hypothesis 1: Given that measures of reaction time and prominence are both related to detectability, a correlation between these two measures is predicted.

This correlation should be a negative one; the more prominent the error, the lower the reaction time.

Hypothesis 2: Since there is no compelling reason for assuming that prominence and importance will correlate, no correlation is predicted.

Hypothesis 3: Likewise, there is no good reason to suppose a correlation between reaction time and importance; none is predicted.

2.4.2. Hypotheses--Stability

Hypothesis 4: As mentioned earlier, it is reasonable to expect that expertise may have a positive effect on error detectionperformance. It is predicted that teachers will be the fastest detectors, beginners the slowest, and intermediates in between.

Hypothesis 5: For the reason given above, we predict that teachers will find errors most prominent, beginners least, and intermediates in between.

Hypothesis 6: Birdsong and Kassen (1988) found no significant between-group differences in relative ratings of error importance; replication of this finding is predicted.

2.4.3. Hypotheses--Error types

Hypothesis 7: Because of the infrequency of English preposition pied piping, and the resulting pervasive difference between English and French behavior with respect to prepositions, preposition stranding errors are hypothesized to take the least time to detect.

Since relativized resumptive pronoun constructions are complex, they are predicted to take the most time to detect. The following order is predicted:

rel. res. pro. > double object dat. > prep. st.

Hypothesis 8: For the same reasons as above, the following ordering is predicted, from most to least prominent:

prep. stranding > double obj. dat. > rel. res. pro.

Hypothesis 9: Given that the relativized resumptive pronoun constructions are ungrammatical in English as well as in French, we predict they will be considered most important. For the same reason stated above, preposition stranding is predicted to be more important. We predict the following importance ordering:

rel. res. pro. > prep. str. > double obj. dat.

These, then, are the hypotheses; we turn now to Chapter 3 for presentation of the study.

2.5. Notes

1. It was felt that reaction times were potentially the most revealing measure of an item's detectability. Although accuracy of response to the error detection task is not used in comparing detectability to prominence and importance

measures, statistical results for accuracy are reported and discussed.

2. "Separating out" is not to be construed as a suggestion that the error exists meaningfully on its own without the context of the sentence. However, for these sentences, it is possible to tell which parts of the sentence are grammatical and which are not.

3. Birdsong and Kassen employed the term "seriousness" rather than the present term of "importance". Nevertheless, the definitions given to subjects in their study and in this one both refer to "importance to correct."

4. Although Nation and McLaughlin use the term "expert" to refer to a learner's experience with several other languages, it should be noted that, used in the context of SLA, the term could also relate to the amount of experience the learner has had with the target language, the learner's metalinguistic skill, the aptitude the learner demonstrates toward the target language, or affective qualities possessed by the learner (e.g. motivation). Investigators such as Naiman et al. (1975), Bialystok and Ryan (1985), Nation and McLaughlin (1986), Schneiderman and Desmarais (1988), Birdsong (1989) and Mclaughlin (1990) have examined each of these areas. With regard to learners in the current study, issues such as experience with the target language, experience with other languages, and metalinguistic skill seem particularly relevant. All are discussed in this chapter.

CHAPTER 3
THE STUDIES

3.1. Tasks and Dependent Measures Used in the Experiments

3.1.1. Error Detection

3.1.1.1. Previous Uses of the Error Detection Task

Most error detection studies in linguistic domains have focused on children, with the aim of discovering how well they detect different types of inconsistencies within a text. Typically, tasks have tested children's ability to note errors that render constructions inconsistent with world knowledge (e.g. The boy rode the apple) or which interfere with comprehension (v. Winograd and Johnston, 1982). Only a few studies have addressed children's ability to detect errors of a formal nature in constructions with plausible interpretations (e.g. Boy is eating apples) (Galambos and Hakuta, 1988). These studies indicate that monolingual children are not successful in detection of formal errors until the age of nearly six, but that bilingual children can detect such errors at the age of 4;6. Galambos and Hakuta state that "young monolingual children appear to focus on the message conveyed by constructions, whereas bilingual children readily focus on the form of constructions upon demand" (p.147).

Error detection studies involving adults have constituted one of several enterprises aimed at discovering

the learner's level of achievement in learning a second language. Four types of grammaticality judgment tasks have typically been employed in second language acquisition (SLA) studies to assess learners' L2 competence. These tasks are the following: 1) discrimination between well-formed and deviant sentences; 2), error detection (or "error identification"); 3) correction of errors, and 4) provision of grammatical descriptions of the errors. (For a review and description of the various studies, see Chaudron, 1983; Ellis, 1991.) It should be noted that error correction and description tasks may or may not involve prior completion of the error detection task, depending on whether or not the error has already been identified for the subject by the researcher.

Of these four task types, tasks which require the learner to discriminate between well-formed and deviant sentences (i.e. to decide grammaticality, acceptability, etc.) are by far the most frequently employed in SLA research. Such tasks often elicit learner intuitions in an attempt to ascertain how well they match those of native speakers (e.g. Mazurkewich, 1985; Tanaka, 1987; Bley-Vroman, Felix, and Ioup, 1988; Cook, forthcoming, etc). The assumption is that learner grammaticality intuitions reflect competence in the L2. The flaw in this claim lies in the neglect of performance constraints involved in the task. This problem is discussed further below in the context of caveats concerning the interpretation of data from (any of

the four types of) grammaticality judgments. The present research does not seek learner intuitions; discussion will therefore focus on the error detection task.

Error detection tasks are significantly less well represented in the SLA literature than are discrimination tasks. In those studies which do employ error detection tasks, this measure usually occurrs in conjunction with that of error correction (e.g. Cohen and Robbins, 1976; White, 1977; Liceras, 1983; Hulstijn, 1984). While some studies present learners' own errors (Cohen and Robbins, White), others (including the present investigation) employ investigator-generated errors.

3.1.1.2. Distinctions Between Error Detection and Well-Formedness tasks

It should be noted that although a distinction is often not drawn in the literature, tasks involving error detection, correction and description differ from a task which requires learners to discriminate between well- and ill-formed sentences. According to Ellis (1991), the former tasks differ from the latter in at least four ways: 1) Intuition is not sufficient in order to detect, correct or describe an error. Ellis points out that these tasks require a response based on "some degree of conscious analysis" (p.162). In discrimination tasks, on the other hand, learner intuition is often what is sought. 2) The detection, correction or description of an error requires, for practical reasons, that the error be local rather than

global in nature. In contrast, a sentence for which an intuition is required may, if deviant, be deviant "everywhere". 3) It seems usually to be the case (probably as a result of #2) that sentences which possess a particular error to be detected (or corrected, described, or explained), are less likely to cause processing problems than are sentences for which intuitions are elicited. In other words, sentences which are deviant in a local manner are not as likely to be difficult to interpret as are those sentences which are deviant in a global manner. 4) Error detection, correction, and description are objective judgment tasks in the sense that there is usually one response identifiable as the correct response (see section 3.2). While the same can frequently be said of discrimination tasks, it is also sometimes the case that even native speakers disagree about the grammaticality or ungrammaticality of a given sentence (see Bley-Vroman, Felix, and Ioup, 1988; Ellis, 1991).

3.1.1.3. Problems With Interpreting Error Detection Data

We have seen that error detection is one type of grammaticality judgment task; as such, it is subject not only to the difficulties discussed in this section, but also to the caveats applied to interpretation of metalinguistic data in general (see section 2.2.3.1). The error detection task is an objective judgment task.

Birdsong (1989) notes that some people are better error detectors than others. This is not the only problem associated with the use of error detection to provide answers to questions about learner competence. Other problems are relevant to each of the four types of grammaticality judgments. Chief among these problems is that none of the tasks is demonstrably free of performance constraints (Birdsong, 1989; Ellis, 1991). Learners may guess if unsure (e.g. if asked for their intuitions), lose patience, or attempt to reject the same number of sentences as they accept. They may ignore an error if they do not know how to correct it. These (and other) behaviors often make interpretation of data difficult. Carroll, Bever, and Pollack (1981) state that learners' intuitions are behaviorally complex, and this observation holds for other types of grammaticality judgments as well. In addition, it is unclear what learners base their responses on. Schachter (1989) notes that subjects may reject sentences not because they are ungrammatical, but because they are difficult to process (see distinction #2, above). This observation applies to native speakers as well as learners, but undoubtedly especially to the latter group.

The error detection task is employed in Experiments 1 and 2; it is the first task to be completed by subjects in Experiment 1, and the second task in Experiment 2, after completion of the Group Embedded Figures Test (GEFT; see Chapter 7). Data from the error detection task are

correlated with other dependent measures in order to address the issue of whether salience is a unitary concept.

3.1.1.4. Use of Reaction Time in Error Detection Tasks

Error detection studies in linguistics are relatively rare and have not, generally, included measurement of reaction times. However, use of reaction time as the major dependent variable is prevalent in the psychology literature on topics such as sensory coding and selective attention, the retrieval of information from long- and short-term memory, parallel and serial information processing, etc. (Pachella, 1974). Response latency is usually defined operationally as the interval between the onset of a stimulus presentation and the initiation of the subject's response.

Although some SLA grammaticality judgment studies have imposed time limits on subject responses (e.g. Hulstijn, 1984), Cook's (1990; forthcoming, cited in Ellis, 1991) studies are, to the best of my knowledge, the only SLA studies which employ the reaction time measure. His (1990) study investigated learner comprehension of reflexive anaphors. In a forthcoming study, Cook examines learner intuitions for whether or not a sentence is correct ("right" or "wrong"). His studies thus do not employ the error detection task. As far as I know, the current studies constitute the only SLA research which measures subjects' reaction times for error detection.

3.1.1.5. The Present Error Detection Task

This task requires subjects to decide, for each of 40 French sentences, whether an error is present. They are instructed to press the appropriate key as quickly as possible without sacrificing accuracy. Although accuracy will be determined and reported, the dependent variable measured in investigation of the research hypotheses will be reaction time (see Chapter 2).

Note that the error detection task is objective in nature, since it is possible to identify a single correct response for each sentence. Mention was made in section 2.2.3.1 of a distinction between objective and subjective metalinguistic tasks, referred to by Chaudron (1983) as judgments and sentiments, respectively.

3.1.2. Error Verification

This task requires subjects to look at a flawed sentence, together with its corrected version beneath it, and state whether the error corrected in the second sentence appeared to be the same as the error they had found in the detection task. Directions (presented in both oral and written forms) stressed that it was not necessary that the correction be identical to what the subjects themselves might have written if asked to correct the sentence, since there are often several ways of correcting a sentence

containing an error. Instead, they were instructed to simply say whether or not the same error had been found.

The error verification task addresses a problem identified by Kellerman (1985) and Birdsong (1989) with regard to error detection: How can the investigator be certain that a judgment reflects identification of just that deviance which the researcher has in mind? Thus, although data from this task are not compared statistically to those from the other three tasks, they are useful for ascertaining that subjects were in fact responding to the error under discussion.

The argument can be made that a more fool-proof method of determining which element a subject is responding to is to have the subject correct the sentence. This format was not employed, for two practical reasons. The first is that the task is not a pen-and-paper task, and accurate type-written corrections are more difficult (and more frustrating, in a second language) on a computer than they are by hand. The second reason involves the practical matter of successfully recruiting volunteers to participate. The experiment was already relatively lengthy; typed corrections might have resulted in hour-long sessions.

3.1.3. Judgment of Error Prominence

To the best of my knowledge, no studies have asked subjects (especially learners) to decide how prominent or noticeable an error seems to them. In this research,

subjects were asked to decide how much an error "stood out" to them (see section 3.3.1.2 for details). In contrast to the error detection task, a decision as to how prominent an error was required a subjective judgment (see "error detection", above). The fact that it was a subjective task was pointed out to subjects, since participants often feel concerned about whether they are performing correctly.

3.1.4. Judgment of Error Importance

Studies of the relative seriousness or importance of an error have typically involved subjects other than learners, with the exception of the Birdsong and Kassen study (see section 1.1). The current studies test both learners and instructors. Since an error can be considered serious or important in a variety of different ways, subjects were given instructions designed to promote a relatively constant interpretation of the task and a realistic context in which to make their judgments. The instructions were adapted both from the original ones presented to subjects by Birdsong, Thoren, and Price, and from instructions given in a subsequent study by Birdsong and Kassen. (For a script of the instructions, see section 3.2.1.2.)

Note that this task, like the prominence judgment task, required a subjective judgment.

3.1.5. The Group Embedded Figures Test (GEFT)

The GEFT is a modified version of the Embedded Figures Test (EFT). The former is designed to be group administered. The GEFT has been shown to be a reliable substitute for the EFT (cf Witkin, Oltman, Raskin, and Karp, 1971, for a discussion). Because of its ease of administration and its low intimidation factor, the GEFT was used instead of the EFT in Experiment 2, even though subjects were tested singly. In this section, we will therefore discuss the GEFT, with the understanding that the discussion is relevant, also, to the EFT.

The rationale for using the GEFT comes from a "perceptual style" or "cognitive style" approach. In this approach, individuals are classified on a continuum ranging from field-dependent to field-independent. Classification is based on the individual's ability to disembed a simple figure (e.g. a triangle) from within a larger, more complex figure (e.g. a geometric design). This classification is not a dichotomous one; the better an individual is at disembedding, the more field independent he or she is considered to be. In other words, the field-dependent/independent distinction is based on the extent to which the parts of a field can be perceived within the organization of the prevailing field.

The GEFT has been found to be a reliable assessor of an individual's ability to disembed. Unsurprisingly, therefore, performance on the GEFT has been found to relate

to performance on a wide variety of tasks which share the requirement of perceptual disembedding, including classical tasks involving illusions and reversible perspective (see Gardner, 1961, and Newbigging, 1954, cited in Witkin et al., 1971). Performance on the GEFT does not, however, relate to performance in perceptual situations lacking this disembedding requirement. In other words, performance on the GEFT relates to tasks which require not simply discriminating parts of a field from each other or discriminating between conflicting stimuli, but rather the "breaking up" of an organized field in order to separate out a part of it (Witkin et al, 1971).

Error detection is a task which requires separating out a part from the whole. Therefore, it is reasonable to expect that ability to detect errors might correlate with score on the GEFT. In other words, the field-dependence/independence distinction may provide a means of characterizing the underlying cognitive processes involved in error detection and relating them to salience. As mentioned previously (see Chapters 1 and 2), a common assumption in the literature is that a salient item is salient to all perceivers and that salience is not susceptible to individual differences. By exploring correlations between GEFT scores and the three present dependent measures (i.e. reaction time and judgments of prominence and importance), we may find a dimension of salience which is traceable to individuals' detection

ability rather than to the inherent detectability of a given error. There is some evidence for a link between field independence and metalinguistic performance. As reported in Kess and Hoppe (1983), Lefever and Ehri (1976) found a positive relationship between field independence and sentence disambiguation abilities. Although Hunt (1985) found no significant correlation between measures of field independence and performance on a variety of language tasks, his study did not include error detection. Present use of the GEFT in Experiment 2 is intended to fill that gap. A sample GEFT item is found in Appendix 3.

3.2. Restatement of the Research Hypotheses

In Chapter 2, we generated a series of hypotheses about what the results of the study might be. These hypotheses are reiterated below:

3.2.1. Unitariness Hypotheses

Hypothesis 1: A negative correlation will be found between Reaction Time and Prominence.

Hypothesis 2: No correlation will exist between Prominence and Importance.

Hypothesis 3: No correlation will exist between Importance and Reaction Time.

3.2.2. Stability Hypotheses

Hypothesis 4: Significant differences will be found between groups on the Reaction Time measure (Teachers lowest, beginners highest).

Hypothesis 5 Significant differences will be found between groups on the Prominence measure (Teachers highest, beginners lowest).

Hypothesis 6: No between-group differences will be found for the measure of Importance.

3.2.3. Hypotheses Regarding Salience of Error Types

Hypothesis 7: With regard to Reaction Time, the following ordering of types is predicted (from greatest to least RT):

rel. res. pro > double obj. dat > prep. stranding

Hypothesis 8: For prominence ratings, the following ordering is predicted (from most to least prominent):

prep. stranding > double obj. dat. > rel. res. pro.

Hypothesis 9: For importance ratings, from most to least important, the following ordering is predicted:

rel. res. pro. > prep. stranding > double obj. dat.

We now examine the study.

3.3. Experiment 1

3.3.1. Method

3.3.1.1. Subjects

Fifty-six individuals participated in the study. Subjects were students and instructors of French from the University of Texas and the University of Florida. There were 15 beginning students, 27 intermediate students, and 14 instructors. The two genders were relatively equally represented in each of the groups, although slightly more than half were female. Beginning students were defined as those who had completed one year of university French (i.e. had just begun third semester French). Intermediate students had completed two years of university French (i.e. had just begun fifth semester French). Instructors were non-native speaker teachers of university French. Intermediate student subjects received homework credit for their participation; all other subjects were volunteers.

Students who had studied languages in addition to French were not excluded from the study. However, all students filled out a questionnaire intended to inform the researcher as to linguistic experience (see section 3.3.1.2). All testing was anonymous, but questionnaires were linked to computer data by means of an identification code.

3.3.1.2. Materials

Questionnaire.

All subjects in the two student groups filled out a survey about their language background, grades received in previous courses, how they rank themselves compared to other students, etc. (See Appendix 1). This was an anonymous questionnaire, which was linked to data by means of an identification code. Instructors did not complete the questionnaire.

Stimuli.

The stimuli consisted of 40 French sentences. Of these, 10 were correct and were distractor items. The other 30 contained grammar errors. No errors of a proof-reading nature were included. These 30 flawed sentences were presented an additional three times: Once, with a corrected version of the sentence underneath, for the error verification task, and twice more, for judgments of importance and prominence. Therefore, each subject viewed a total of 160 French sentences.

Errors in the sentences were of three basic types: Preposition stranding, double object dative constructions, and relativized resumptive pronoun constructions. Subjects were instructed that no proof-reading errors (e.g. missing or wrong accents) occurred. The following are examples of each type of flawed sentence (see Appendix 2 for a complete list):

Prep. stranding - question *Qui est-ce que Luc voit le film avec? *(Who is Luc seeing the movie with?)*

Prep. stranding - rel. clause *Voila l'appartement que Thomas habite dans.*(Here is the apartment that Thomas lives in.)*

Double object dative construction *Paul donne Suzanne un livre. *(Paul is giving Suzanne a book.)*

Rel. res. pro. constr. - subject *La femme qui elle est dans la rue est blonde. *(*The woman who she is in the street is blond.)*

Rel. res. pro. constr. - dir. object *La femme que Luc la cherche dans la rue est blonde. *(*The woman who Luc is looking for her in the street is blond.)*

Rel. res. pro. constr. - indir. obj. *La femme que Jean lui parle est sympa. *(*The woman who John is talking to her is nice.)*

The 10 correct sentences were not correct versions of any of the deviant sentences, but rather sentences with structures which bore a superficial similarity to those items being tested. (e.g. "Qui est dans la voiture avec ta soeur?" is superficially similar to "Qui est-ce que Luc voit le film avec?".)

Apparatus.

Stimuli were presented and responses taken on an IBM-compatible personal computer. All relevant keys were labelled. The computer did not respond to stray key touches.

Tasks

A brief enumeration of the tasks is now provided for the reader's convenience. For a more detailed discussion, see section 3.1.

1. *Error detection.* This was the first task performed. Subjects viewed 40 French sentences one by one on the computer monitor. For each sentence, they were asked to decide as quickly as possible whether or not the sentence contained an error. They were instructed to press a key labelled "yes" if they found an error. They were asked to hit the "no" key if they found no error.. Subjects were told not to look for errors of a proof-reading nature, but to concentrate on grammar errors. They were instructed to respond "as quickly as possible without sacrificing accuracy". Response times were recorded. Before beginning the test, subjects completed four practice items.

Of the 40 sentences, 30 contained errors. That some sentences were correct was not explicitly pointed out to subjects. Nevertheless, they were told the task involved

making a decision; furthermore, two of the four practice items were correct sentences. The purpose of the 10 correct distractor sentences was to keep subjects on task.

2. *Error verification.* Of the 40 sentences, 30 contained a grammatical error. After completion of the error detection task, each of the 30 flawed sentences was again presented, this time with the corrected version shown just below it. Subjects were asked to indicate whether the error which had been corrected was the same error as they themselves had found. If it was, they pressed "yes". If it was not, or if they had not found any error for that sentence, they were to hit the "no" key. It was pointed out to subjects that there may exist more than one correct version of a given sentence. Instructions emphasized that the task was to indicate whether the author of the correction "seems to have found the same error you did, not whether they corrected it the way you would have".

3. *Judgment of error prominence.* The 30 flawed sentences were presented and subjects gave each error a rating as to how "glaring" it appeared, on a scale from 1 (not at all glaring) to 6 (very glaring). The scale and its end labels remained on the screen for each of the sentences, to eliminate the possibility of any confusion. Subjects were given the following instructions:

You are asked to view each sentence and say how much the error in it stands out to you; in other words, you are asked to judge how "glaring" it is. Note that this is a purely subjective judgment, for which there can be no "correct" or "incorrect" answer.

These instructions were given to subjects both orally and in writing, and in each case appeared after the subject had completed the error detection and verification tasks. For half of the subjects, the judgment just described took place before the judgment of error importance (see below). For the other half, judgment of error importance was performed before judgment of error prominence. These subjects received the following additional directions:

Note that judging how glaring an error is is not the same task as judging its importance; your answers may or may not be similar.

These directions were worded in such a manner as to encourage subjects to approach the task relatively uninfluenced by the task just completed, without giving the impression that any particular sameness or distinctness of responses was sought by the investigator. In other words, the goal was to keep the two concepts of prominence and importance distinct in the minds of subjects without leading

them to making a distinction that was artificial or exaggerated.

4. *Judgment of error importance*. The 30 flawed sentences were once again presented. This time, subjects gave each error a rating as to how important it was to correct. The scale ranged from 1 (not at all important) to 6 (very important). The term "important" was not narrowly defined; subjects were given the following scenario for use as a guide:

Imagine that you have been helping your friend, who knows you study (teach) French, to improve her own French in preparation for a trip to France. You have written down errors made by her in casual conversation with you. Assuming that the more serious an error is, the more important it is to correct it, and that it is not possible to correct all her errors, how important is it that each of the following errors be corrected?

As with the judgment of error prominence, the scale and its end markings remained on the screen. Half of the subjects performed this task before making prominence judgments (see above); for the other half, the judgment of importance was the fourth and final task. These subjects received additional instructions (see above rationale):

Note that judging how important an error is is not the same task as judging its prominence; your answers may or may not be similar.

3.3.1.3. Procedure

All subjects in the two student groups first filled out a questionnaire about their language background. After this procedure (which was not carried out for the instructor group), subjects were seated in front of PC computers for testing. The four tasks (error detection, error verification, judgment of error prominence, and judgment of error importance) were performed by all subjects.

Occasionally, more than one subject at a time was tested. However, in these instances, subjects had their own computers and their own separate environment (small computer rooms). Instructions were all in English. They appeared on the screen and were read to each subject. Any additional directions required were given orally, and instructions were repeated as many times as needed for complete comprehension of the task. Subjects were instructed to inform the investigator when they reached the end of a section; in this manner, oral as well as visual directions were provided. To maximize subjects' general comfort, the investigator left the immediate vicinity during performance of each task, while remaining within earshot.

Sentences remained on the screen until a response was made, but subjects could not go back and review sentences.

Only the error detection task was timed; subjects were told they could take their time on all other tasks.

Attempts were made to control for some of the factors which have been suggested to affect salience (see section 1.2). For example, the notion that humanness is of more interest to humans than is nonhumanness (v. Fillmore, 1977) is accommodated by the inclusion of a human in each stimulus sentence. The study also attempts to control for the degree of concreteness (v. Paivio et al., 1968, and others) in structures tested, as much as possible. Thus abstract nouns such as "love" are not employed.

Hatch (1983) states that free morphemes are more perceptually prominent than are bound ones. Since prepositions such as "a" and "de" are often (although not always) bound, these two prepositions have been excluded from the preposition stranding exemplars.

In addition, controls have been made, as much as possible, for length of sentence across all stimulus items and for choice of pronoun (always the feminine) in the relativized resumptive pronoun constructions.

It is hoped that all of these measures preserve the integrity of the data.

3.3.2. Results

In order to determine whether correlations exist between the dependent variables, correlation analysis (Spearman, since we do not have interval data - see endnote

1) was performed, using data for all proficiency groups and all error types combined. As shown in Table 3.1, below, correlation analysis reveals a significant correlation, overall, between Importance and Prominence ($p<.001$). Similarly, a significant correlation is found between Prominence and Reaction Time ($p<.001$). The former is a positive correlation, while the latter is a negative one; in other words, as Importance increases, so does Prominence, but an increase in Prominence correlates with a decrease in Reaction Time. Both of these findings make intuitive sense. Even though we predicted (see Hypothesis 2, section 3.2) that prominence and importance would not correlate, it is nevertheless a common expectation that a relatively more important error will also be relatively more prominent. That this idea is widespread, in fact, explains why discussions of salience in the literature often do not distinguish between aspects of prominence and importance (see Chapter 1). Similarly, it makes sense to suggest that an error which stands out relatively more than another should take less time to detect, hence a high prominence rating should correspond to a low reaction time. The expectation that detectability, as measured here by reaction time, should correlate with subjects' prominence judgments was discussed in section 2.1.5.

So far, then, we find that Hypothesis 1, which predicts a negative correlation between Reaction time and Prominence ratings, is supported by the results, while Hypothesis 2,

predicting no correlation between Prominence and Importance, is not. Both of these findings suggest unitariness (predicted in Hypothesis 1 but not in Hypothesis 2). However, consistent with the prediction of Hypothesis 3, no correlation is found between Reaction time and Importance. Thus we find a conflicting suggestion of a lack of unitariness. This result may be partially explained by considering the argumentation presented above and in section 2.1.5. It was mentioned that Reaction Time and Prominence ratings are both a function of detectability, while Importance is not. And although Prominence and Importance judgments might conceivably be expected to covary, especially since both are subjective judgments, no such argumentation can be found to suggest that Importance judgments should correlate with Reaction Time.

Although the correlations between Importance and Prominence and between Prominence and Reaction time are significant at the $p<.001$ level, neither of them is a particularly strong one since, in each case, absolute values of the coefficients are quite far from 1. What this apparent contradiction implies is that the correlations are relatively weak but persistent. This situation is not an unusual one; Kincaid (personal communication) points out that significance levels of $p<.001$ are often found in studies with large numbers of data points. He further suggests that in such situations, statistical results of $p<.001$ should be taken "with a grain of salt".

Table 3.1. Spearman Correlation Coefficients

Comparison			Spearman Correlation Coefficient
Importance	vs.	Prominence	+.369***
Prominence	vs.	Reaction time	-.164***
Reaction time	vs.	Importance	-.018

Note: "***" indicates p<.001

A Multivariate ANOVA was next performed. This was a 3 X 6 factorial ANOVA with the independent variables GROUP (beginning, intermediate or teacher) and ERROR TYPE (1,2,3,4,5 or 6 - see Table 3.3.) and dependent variables IMPORTANCE, PROMINENCE, and RT (reaction time).

The ANOVA shows overall effects (i.e. over all tasks) for proficiency level (hereafter, GROUP), the type of error involved (ERROR TYPE) and the interaction between the two. As can be seen in table 3.2, effects for GROUP and ERROR TYPE both achieve the p<.001 level of significance, while the interaction GROUP X ERROR TYPE is significant at the p<.01 level.

Table 3.2. ANOVA - Effects over all tasks

	F-value	p-value
GROUP	55.6831	.0001***
ERROR TYPE	20.7313	.0001***
GROUP X ERROR TYPE	1.9441	.0015**

During the remainder of this dissertation, particularly in the diagrams, ERROR TYPE will be referred to as a number, for simplicity. Table 3.3 provides the key.

Table 3.3. Explanation of number references to ERROR TYPE

Error Types

1= Preposition stranding, question
2= Preposition stranding, relative clause
3= Double object dative construction
4= Relativized resumptive pronoun - subject
5= Relativized resumptive pronoun - direct object
6= Relativized resumptive pronoun - indirect obj.

The next step is to consider each of the three proposed salience components in turn. We first examine the measure "Importance".

3.3.2.1. Importance

Results of analysis of the dependent variable Importance are shown in Table 3.4.

Table 3.4. General Linear Models Procedure - Importance

	F-value	p-value
GROUP	1.45	.2349
ERROR TYPE	26.22	.0001***
GROUP X ERROR TYPE	1.64	.0902

Significant effects are found only for ERROR TYPE (F= 26.22; p <.001). Therefore, only ERROR TYPE undergoes any further analysis. Figure 3.1 shows the results of the Tukey's Studentized Range test for the dependent variable Importance with regard to ERROR TYPE. Rankings shown are for all groups combined.

Note: "*" indicates significance value of p<.05.

Figure 3.1. Tukey's Studentized Range test: IMPORTANCE (all groups combined)

As can be seen from Fig. 3.1, the farther one goes to the left end of the scale, the more important a given error type was judged to be, so that the positioning of an error

type to the left of another type indicates that the former was judged more important than the latter. Using the key in Figure 3.3 to decipher error type, we see that the relativized resumptive pronoun constructions (i.e. types 4-6) were judged most important. The two kinds of preposition stranding (1 and 2) were found next most important, and the double object dative construction was judged least important.

Asterisks indicate statistical significance. Thus we see that the three relativized resumptive pronoun constructions were judged significantly more important than the other three error types ($p<.05$). The preposition stranding-question error type (type 1) was found significantly more important than the double object dative construction (type 3). As mentioned above, these importance ratings are overall ratings; they are not broken down by GROUP, since between-group differences were not found to be significant. Recall that Hypothesis 9 predicted that importance ratings would be in the following order, from most to least important:

rel. res. pro. > Prep. stranding > double obj. dat.

Thus Hypothesis 9 is found to be correct.

As mentioned earlier, the only effect found for IMPORTANCE was for ERROR TYPE; no significant differences were found between the subject groups with respect to judgments of importance for the structural variables. Nevertheless, the means and standard deviations may be of

interest; these are displayed, for each subject group and each error type, in Table 3.5, below.

Table 3.5. Means and standard deviations by group and error type: IMPORTANCE

		1	2	3	4	5	6
Beginner	M	3.8	3.9	3.5	4.4	4.2	4.1
	SD	1.6	1.6	1.9	1.6	1.4	1.3
Intermed.	M	3.9	3.7	3.5	4.4	4.2	4.5
	SD	1.2	1.3	1.5	1.3	1.3	1.1
Instruct.	M	3.9	3.6	3.3	4.7	4.7	4.8
	SD	1.3	1.4	1.6	1.5	1.4	1.3

Note that although it looks as if some of these differences, especially between beginners and teachers, ought to be significant, standard deviations are considerably higher than any of the differences.

We do not find significant inter-group differences in ratings for Importance, in either absolute or relative terms. Thus Hypothesis 6 turns out to be correct, and we have some (partial) support for the notion that salience is a stable concept.

3.3.2.2. Prominence

Analysis of the dependent variable Prominence reveals more effects than were found for Importance. There are significant ($p<.001$) effects for GROUP, ERROR TYPE and the

interaction GROUP X ERROR TYPE. These are shown in Table 3.6, below.

Table 3.6. General Linear Models Procedure - PROMINENCE

	F-value	p-value
GROUP	26.35	.0001
ERROR TYPE	39.27	.0001
GROUP X ERROR TYPE	3.30	.0003

Hypothesis 5 predicted significant between-group differences; these were found.

Table 3.7, below, shows means and standard deviations of Prominence for each subject group response to each type of error.

Table 3.7. Means and standard deviations by Group and Error type: PROMINENCE

		1	2	3	4	5	6
Beginner	M	3.9	4.0	2.8	4.6	4.1	3.9
	SD	1.7	1.6	1.5	1.6	1.3	1.4
Intermed.	M	4.6	4.5	3.7	5.3	4.5	4.3
	SD	1.5	1.6	1.8	0.9	1.3	1.5
Instruct.	M	3.8	4.0	3.6	5.3	5.0	5.0
	SD	1.7	1.6	1.7	1.0	0.9	1.0

Figure 3.2 gives the results of the Tukey's Studentized Range test for the dependent variable "Prominence", broken

down by group. Note that the further to the left the positioning of a given error type, the more prominent it was judged to be. Immediately obvious is the fact that error types at extremes of the range (i.e. types 3 and 4) are the same for all groups; in other words, each of the groups judged relativized resumptive pronoun constructions-subject most prominent and the double object dative construction least prominent.

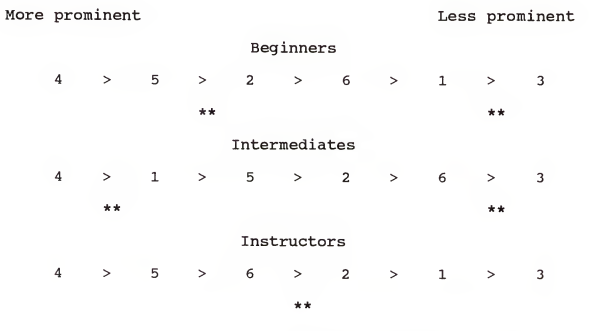


Figure 3.2. Tukey's Studentized Range test: PROMINENCE

The arrangement of asterisks in the diagram above indicates that, for both beginners and intermediates, error type 4 was rated significantly more prominent than all other error types and type 3 was judged significantly less prominent than all other types. For instructors, types 4,

5, and 6 (i.e. relativized resumptive clause constructions) were all judged significantly more prominent than each of types 2, 1, and 3. As noted in several of the figures above, the presence of two asterisks indicates a $p<.01$ level of significance.

Recall that Hypothesis 8 predicted the following error type ordering for ratings of prominence (from most to least):

prep. stranding > double obj. dat > rel. res. pro.

Hypothesis 8 is not supported by these results. Of interest is the fact that prominence rankings of error types are the same for instructors and beginners, except for a reverse ordering of types 2 and 5 at the middle of the range. Similarly, beginner and intermediate prominence rankings differ only in the placement of error type 1 (preposition stranding); all other types follow the same ordering. Intermediate and instructor orderings differ with regard to error types 1 and 6.

Figure 3.3 is a plot of the mean Prominence ratings given to each of the error types by the three different subject groups. This plot was generated by analysis of Least Significant Differences (LSD's) between means. As the key shows, filled squares indicate beginners, empty squares are intermediates, and X's indicate instructors. A first glance at the plot reveals that in general, beginners find the errors to be less prominent than do either intermediates or instructors.

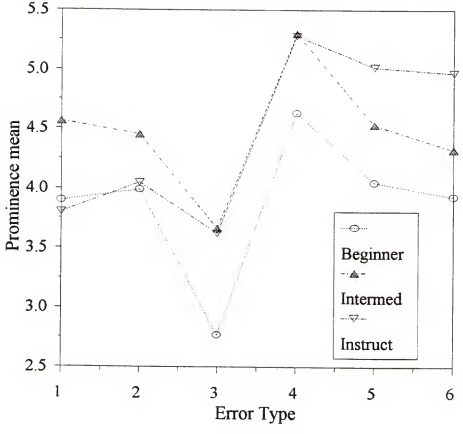


Figure 3.3. Plot of prominence means for error types

In spite of the fact that beginners and instructors have (visually) very similar relative prominence ratings of error types, as mentioned above, examination of means reveals that the only error types for which absolute prominence ratings (that is, the actual numbers) for beginners and teachers are similar in magnitude are types 1 and 2 (i.e. preposition stranding types). Intermediate judgments for these types are quite different (although significantly so only in the case of type 1). Instructor means for all non-preposition stranding error types are considerably higher than beginner means, and for types 3 and 4, intermediate means closely match instructor means.

To avoid confusion, asterisks indicating levels of significance have not been shown on the Fig. 3.3 plot; these will now be discussed. Notice that two out of the three means cluster together (in most cases, quite tightly) for every error type except type 5, where the three group means are evenly spaced apart. For type 5, beginner and instructor means are significantly different from each other ($p<.01$), but are not significantly different from the intermediate mean. For each of the other error types (with the exception of error type 2, for which there are no significant differences), the "marooned" mean is significantly different from the other two, at the $p<.01$ level.

3.3.2.3. Reaction Time

Further analysis of the dependent variable Reaction Time reveals a main effect for GROUP and also for ERROR TYPE; as shown in Table 3.8, both achieve a $p<.001$ degree of significance.

Table 3.8. General Linear Models Procedure - Reaction Time

	F-value	p-value
GROUP	152.37	.0001
ERROR TYPE	9.61	.0001
GROUP X ERROR TYPE	1.22	.2738

Data analyzed to produce table 3.8, above, do not include reaction times for grammatical sentences. However, the latter are very similar to reaction times for ungrammatical sentences.

Table 3.9, below, shows mean RTs and differences between RTs for each of the groups (i.e. by GROUP).

Table 3.9. Mean RTs and differences between RTs by GROUP -errors combined

(SLOWER)	Beginners	Intermed.	Instructors	(FASTER)
	10.99 sec	8.48 sec	4.39 sec	

Unsurprisingly, beginners are at the slow end of the scale, with a mean reaction time of 10.99 sec. We see that

they are 2.5 seconds slower than the intermediates, at a significance level of $p<.05$. Intermediates are, in turn, significantly ($p<.05$) slower than instructors. Note that the Table 3.9 data are for all error types combined.

Table 3.10 shows mean reaction times for each of the six error types, and indicates significant differences between means, where they exist.

Table 3.10. Mean RTs and comparisons between RTs for each ERROR TYPE

(SLOWER) ERROR TYPE	6	5	1	2	4	3 (FASTER)
	9.76	8.23	8.34	7.63	6.98	6.83
6						
5						
1	*					
2	*					
4	*	*	*			
3	*	*	*			

As can be seen in Table 3.10, reaction times for error type 6 are significantly different, $p<.05$, from those for every other error type except 5. Error types 3 and 4 have reaction times which are significantly different, $p<.05$, from those of all error types (not including each other) except 2. Note that Table 3.10 combines all subject groups.

Table 3.11, below, shows means and standard deviations of REACTION TIME for each subject group response to each type of error.

Table 3.11. Means and standard deviations by group and error type: REACTION TIME

		1	2	3	4	5	6
Beginner	M	11.4	11.2	9.0	9.7	11.4	13.4
	SD	7.9	8.0	5.4	6.2	6.3	9.2
Intermed.	M	9.0	7.8	7.0	7.6	9.4	10.1
	SD	4.8	4.7	3.6	4.5	6.3	6.6
Instruct.	M	4.6	4.0	4.5	3.7	3.8	5.8
	SD	2.8	1.9	3.2	2.6	2.1	4.7

A glance at mean reaction times in Table 3.11 reveals quite slow reaction times for beginners. Note, however, that the standard deviations are also quite large. In fact, the beginner data included several reaction times of over a minute, which accounts in part for the figures shown. The same situation, although to a much lesser extent, holds for the intermediates. Since some beginner reactions times were so long that they could conceivably have distorted the means, median reaction times were also calculated for beginners. However, in most cases the median was very close to the mean.

Hypothesis 4 predicted that significant inter-group differences would be found for Reaction Time, and that reaction times for teachers would be lowest (i.e. response

fastest) and those for beginners would be highest (i.e. response slowest). Thus Hypothesis 4 is shown to be correct.

3.3.2.4. Other results

3.3.2.4.1. Accuracy and verified accuracy

Accuracy is defined in terms of the percent correct for each group on the error detection task. This percentage includes "hits" (correct answers of "yes") and correct rejections (correct answers of "no error"). Verified accuracy is the percentage of "hits" which received answers of "yes" on the verified accuracy task. Table 3.12 shows these results; additional explanation is given below.

Table 3.12. Accuracy and verified accuracy (%) by GROUP

	Accuracy	Verified accuracy
Beginners	72.65	80.89
Intermediate	78.07	89.97
Instructors	94.35	96.66

Chi-Square tests reveal significant differences for accuracy between all groups. Beginners were significantly less accurate than both intermediates and instructors, and intermediates were significantly less accurate than instructors. All findings are at the $p<.05$ level.

Verified accuracy measures the percent of the time that the locus of a correctly identified error was said to be the

same as that shown in the corrected version. In other words, it measures how often, for all correctly identified errors, the subject indicated that the error (s)he found matched that which was shown corrected in the error verification task. Again, all inter-group differences were statistically significant ($p<.001$).

These results reveal that instructors also had some difficulty with the error detection task, since accuracy percentages are not 100%. The most likely explanation for this imperfect accuracy is that errors were made in hitting the correct key. Since the accuracy percentages are problematic, no further Experiment 1 analyses were made using accuracy. A more detailed discussion of accuracy and verified accuracy scores appears in section 3.4; additional discussion in undertaken in Chapter 5.

3.3.2.4.2. The questionnaires

It will be recalled that all learners (but not instructors) filled out a questionnaire intended to inform the researcher as to their amount of exposure to French as well as to other languages (see Appendix 2). Based on these questionnaires, it is difficult to separate factors of exposure and expertise. Of the 15 beginners, 12 had had at least a year's experience with a foreign language other than French. For the intermediate group, 24 out of 27 had had such experience. Two-thirds of the beginners received

either a grade of A or B; only 3 intermediates had grades of C or below.

The difficulty of defining novices and experts has been mentioned in section 2.2.2. We will take up this issue again in Chapter 5.

3.3.3. Discussion

A global discussion of the results of this experiment, as well as those of Experiment 2, within the context of the questions posed by this research, is found in Chapter 4. Present discussion will focus on some comments which need to be made regarding Experiment 1.

As we have seen, the correlations found between Importance and Prominence and between Prominence and Reaction Time, although highly significant, are nevertheless rather weak (see section 3.3.2 for a discussion). An additional problem is that the two subjective judgments of Importance and Prominence were made contiguously in all cases (although in two different orders). It is therefore possible that a blurring of definitions may have occurred.

Main effects for GROUP and ERROR TYPE exist for the dependent variables Prominence and Reaction Time, and, in addition, an interaction GROUP X ERROR TYPE is found for Prominence. Of the three dependent variables, that of Importance exhibits the fewest effects; only ERROR TYPE is significant. This lack of effect for proficiency level is

especially interesting because it results in two findings: 1) instructors and students agree on the absolute importance (i.e. magnitude of judgments) of the error types in this experiment and 2) instructors and students agree on the relative importance of error types. Although this first finding runs counter to results obtained by Birdsong and Kassen (1988), the second finding replicates their discovery of agreement among teachers and students regarding relative importance of errors.

How can these findings be explained? It is possible that students gain clues about which errors instructors consider most important, either by their reactions to errors in class or by their grading procedures. In this way, learner judgments of error importance may be more likely to become "matched" to those of instructors. For more discussion of likely reasons, see section 2.2.2.

That ERROR TYPE has a significant effect for each of the three dependent measures is an interesting and consistent result. The situation is somewhat complicated, however, by the fact that error types 4, 5, and 6, all of which are relativized resumptive pronoun constructions, only cluster together on the measure of judgments of importance; with respect to reaction times and judgments of prominence, subject behavior appears inconsistent among the three types. In other words, the fact that types 4, 5, and 6 are all relativized resumptive pronoun constructions does not produce consistent treatment of them by subjects. This

finding provides support for Clark's (1973) caution regarding the language-as-fixed-effect fallacy.

A further complication in interpretation is the fact that, among the six types of errors, those judged most important or most prominent are not necessarily associated with the fastest reaction times. In other words, it is clear that error type distinctions, in the general sense, matter; however, particular types of errors are not ranked in a consistent manner across the three proposed salience components. For example, error type 3 (double object dative construction) was judged least important and least prominent, but is nevertheless associated with the fastest reaction times.

Error types 4 and 5, and, to a lesser degree, 6 (all of which are relativized resumptive pronoun constructions) are judged significantly more important and more prominent than the other error types. Mention should be made again here (see also Chapter 1) of the fact that, unlike the other three error types, all three of the relativized resumptive pronoun constructions are ungrammatical in English as well as in French; any effect this fact may have had on ratings, however, appears to apply mostly to prominence ratings by instructors (see Fig. 3.3).

Interestingly, for each of the two subjective judgment tasks (i.e. importance and prominence), beginners, as a group, used the entire scale range (1-6) when making their judgments. This does not mean that every beginner used the

entire scale; however, the fact that the extremes of the scale (1 and 6) were used is worth noting, since the comment is often made in the SLA literature that novices (beginners) tend to stay in the middle of scale ranges, in general.

In broadest terms, based on the results just described, what we conclude about the three research questions posed in this dissertation is the following: 1) Salience does not appear to be a unitary concept. (The measures of reaction time, judgment of prominence and judgment of importance correlated either weakly or not at all.) 2) Salience appears to be a stable concept only if it is defined as Importance. 3) Some types of errors appear to be more important than others, some are more prominent than others, and some are more detectable. Therefore, which error types are more salient depends on which measure of salience is employed (see above discussion).

More will be said about these results and what they mean for the present research in the General Discussion in Chapter 4. We turn now to Experiment 2.

3.4. Experiment 2

3.4.1. Purpose

The purpose of this experiment is to determine whether there is a correlation between subject performance on two of the tasks administered in Experiment 1 (error detection and error verification) and performance on the Group Embedded Figures Test (GEFT).

The GEFT is a perceptual test of how well an individual can disembed a simple (previously seen) figure from within a larger, more complex figure. Otherwise stated, the test assesses the extent to which the organization of the prevailing field dominates perception of its parts. The rationale for using the EFT (from which the GEFT is derived; see section 3.1.5) comes from evidence that competence at disembedding in perceptual tests is strongly associated with ability to disembed in non-perceptual problem-solving and with performance in a very wide array of perceptual situations which share the requirement of perceptual disembedding (Witkin, Oltman, Raskin, and Karp, 1971).

As mentioned above (see section 3.1.5), there is some support for a link between field independence and metalinguistic performance. In this experiment, we seek to discover whether this psychological correlate (field independence) underlies detectability.

3.4.2. Method

Most of the methodology involved in Experiment 2 is identical to that of Experiment 1 (see section 3.2.1). Methodology concerning the Group Embedded Figures Test (GEFT) is discussed in this section, and brief reminders of the Experiment 1 methodology are provided.

3.4.2.1. Subjects

Eighteen beginning students of French at the University of Florida, 6 males and 12 females, participated in the experiment. (As in Experiment 1, a rank of "beginner" indicates completion of one year of college French.) Each subject received $5 compensation for completion of the tasks. Again, students who had studied additional languages were not excluded from the study; questionnaires were linked to computer data by means of an identification code.

3.4.2.2 Materials

Questionnaire

The questionnaire filled out by subjects prior to performance of any of the tasks was identical to that of Experiment 1; it contained questions regarding language background, grades received, etc. (see Appendix 1).

Stimuli

The stimuli for the first task, completion of the Group Embedded Figures Test (GEFT), consisted of three series of simple figures embedded in a larger complex figure (cf description of the GEFT, below). For the remaining tasks, sentences were identical to those in Experiment 1; once again, 30 of the 40 sentences contained errors.

Apparatus

The GEFT is a timed pencil-and-paper task. A digital hand-held timer was employed. All other tasks were carried

out on an IBM-compatible personal computer, as in the preceding experiment.

Tasks

Tasks were the same as those performed in Experiment 1, except that an additional task, the GEFT, was administered. A more thorough discussion of this paradigm is found in section 3.1.5. In addition, the test is discussed in section 3.4.1 and in this section, below. A brief reminder is provided here as to what the other tasks were; for a more detailed discussion, see section 3.4.1.2.

1. *The Group Embedded Figures Test (GEFT).* As previously stated (see section 3.1.5), the GEFT is a modification of the Embedded Figures Test (EFT). The two tests have been shown to be equivalent with respect to reliability and validity (Witkin et al., 1971); the GEFT was used for ease of administration and maximum subject comfort. Subjects were required to locate a previously seen simple figure within a larger complex figure which had been arrayed in such a manner as to obscure or embed the simple figure. (See Appendix for a sample problem.) Subjects could refer to the simple forms on the back of the booklet as many times as they wished. After finding the simple figure, subjects traced over it in pencil. They were given 2 minutes for the first section and 5 minutes for each of the other two sections. The highest possible score on the GEFT is 18. The first section is not scored; subjects receive 1

point for each figure in sections 2 and 3 which is completely and correctly traced.

2. *Error detection*. Completed after the GEFT, this task was the same as that in the preceding experiment. Once again, subject timing was achieved through a device in the computer software. Subjects were instructed to respond "as quickly as possible without sacrificing accuracy".

3. *Error verification*. As in Experiment 1, subjects were asked to indicate whether the error which had been corrected was the same error as they themselves had found. The same directions were given; therefore, it was stressed in this experiment, also, that it was the finding of the same error which was crucial, not whether the correction was the one the subject would have made.

4. *Judgment of error prominence*. Subjects gave each error a rating as to how "glaring" it appeared, on a scale of 1 (not glaring) to 6 (very glaring).
The task was administered exactly as in Experiment 1, with the scale and its end labels remaining on the screen. Also as in the preceding experiment, half of the subjects performed this task before judging error importance, and the other half completed it afterwards.

5. *Judgment of error importance*. Subjects gave each error a rating as to how important it was that the error be corrected. The same guidelines were given as for experiment 1; the scale was from 1 to 6.

3.4.2.3. Procedure

All participants were tested individually. Subjects first completed the language background questionnaire.

Next, the GEFT was administered. During this task, subjects were seated at a desk with a pencil and the test booklet. The test is composed of three sections, each more difficult than the preceding. The first section contains seven relatively easy problems (i.e. complex figures); the second and third sections are more difficult and each have nine figures. Each problem indicates which of the forms (shown on the back cover of the booklet, to prevent direct comparison) the subject is to find. Subjects may look at the simple figure as many times as they wish. Once subjects have found the simple figure within the larger more complex array, they are required to trace over it in pencil.

Subjects first completed the two practice figures provided in the booklet. They were then given 2 minutes to complete the first section, 5 minutes for the second, and 5 minutes for the final section. Subjects were not permitted to begin a new section until time was up on the previous one. Timing was effected on a hand-held digital timer.

After completion of the GEFT, subjects moved to the computer for the rest of the tasks. As in Experiment 1, the error detection and verification tasks required YES or NO responses; judgments of importance and prominence required

number responses. The error detection task was computer-timed, but reaction times were not taken on the other tasks.

3.4.3. Results and Discussion

Analysis was performed on correlations between each of the following: 1) GEFT score and overall accuracy; 2) GEFT score and verified accuracy (see below for explanation of the distinction between overall and verified accuracy); 3) GEFT score and reaction time on the sentences containing errors, and 4) GEFT score and reaction time on all sentences (i.e. including distractors).

Overall accuracy was defined as percent correct; that is, the percent of responses which correctly identified the presence or non-presence of an error. Among the eighteen subjects, overall accuracy ranged from 98% to 50%. For the group, the mean overall accuracy was 76.4%.

Verified accuracy was calculated using only sentences which contained errors. It will be recalled that only flawed sentences were presented again for verification that the error corrected was indeed the error the subject found. Of the errors correctly detected, the percentage which were verified is 65%, as a mean for the group. This is not an impressive finding, a fact which will be discussed below.

The mean score on the GEFT was 15, out of a possible 18. It should be noted, however, that fully half of all subjects received either a 17 or an 18 on the test; that the

mean is not higher results from three very low scores (one score of 9 and two scores of 8).

None of the analyses performed yielded statistically significant correlations. In other words, in this experiment, we fail to find a psychological correlate of detectability. Data for each subject is in Appendix 4.

The first observation to be made, as in Experiment 1, is that the timed error detection task is a challenging one. In spite of the four practice items and the labels on the keys, many subjects informed the researcher, after the session, that they had occasionally hit the wrong key.

Numerous questions arise as a result of these findings. First of all, why is verified accuracy so low? One possibility is that, in spite of the instructions, subjects did not understand that it was not necessary that they agree with the way in which a particular sentence was corrected, merely that they identify sameness of error. The fact that there were some NO validation responses from the instructors in Experiment 1 lends credence to this suggestion, especially since the sentences most frequently producing a NO response were the relativized resumptive pronoun constructions involving indirect objects (i.e. sentences 26-30), and although the corrections for them do address the error, it is true that the corrections are not the optimal ones. To suggest that the instructors in question did not find the intended errors is not a very plausible interpretation; none of the errors were of an advanced

nature. In fact, all were intended to be easily accessible even by the beginner group.

It may also be that subjects did not remember the error, and thus replied NO when asked to verify. A related possibility is that subjects may have correctly identified an error, but not known how to correct it and therefore responded NO. Still another possibility is that the lower than expected number of YES verification responses reflects subject use of one of the performance strategies discussed in section 5.1. It was pointed out that subjects sometimes attempt to "balance" their types of responses. Note that a verified accuracy of 100% (exhibited by one subject) requires the subject to have responded YES to every item in the verification task.

The possibility also exists, naturally, that subjects simply did not find the errors that they were intended to find, but found other (imagined) errors instead. The pilot study conducted prior to Experiment 1 did not indicate a problem; nevertheless, as mentioned in section 3.1.2, Kellerman (1985) has noted the difficulties involved in ascertaining that learners and investigators are dealing with the same errors.

In view of the current results, the verification task appears to be more problematic than a task which requires subjects to correct the error. Any follow-up study (see Chapter 5) will need to consider this fact.

Questions arise also about the GEFT test. That so many subjects (half) had either a perfect score or missed only one is surprising, given the long and successful history of the test at evaluating subjects' ability to dissociate figure from ground (i.e. to disembed a simple figure from within a more complex figure). Lack of variability due to an apparent ceiling effect in the GEFT scores may also explain the lack of correlations.

Thus each of these problems may play a role in the lack of significant findings for Experiment 2.

3.5. Notes

1. Data are assumed to be ordinal in nature rather than interval, for reasons which are best illustrated by comparing subject ratings in the present study to student scores on a test. For our data, we cannot be sure that ratings are defined in a constant manner across all subjects, i.e. that a rating of "3" means the same thing to everyone. In contrast, a score of 90 on a test means the same thing (i.e. that ten points were missed) regardless of which student receives it. Furthermore, the interval between, say, 60 and 70, is the same as that between 80 and 90. The same cannot be said for the present data, since there is no guarantee that for all subjects, the interval between ratings of 3 and 4 is the same as that between 5 and 6.

2. Terminology has differed over the years (see Witkin et al., 1971).

CHAPTER 4
GENERAL DISCUSSION

4.1. Review and Discussion of the Major Findings

The present studies are very much preliminary ones, so the results attained must be viewed with caution. Since experiment 2 did not yield significant results, discussion of it will be undertaken in Chapter 5, in the context of limitations of the study and directions for further research. We will focus here on the results of Experiment 1.

Three main research questions were investigated in this dissertation. The first was whether or not salience is a unitary concept. As previously discussed, in order to answer this question, it was necessary to discover whether the three proposed salience components (Detectability, Prominence, and Importance) correlate with each other. The present results show correlations in two out of three comparisons. Importance and Detectability do not correlate, however. Furthermore, the two correlations which are significant are nevertheless very far from 1.0 and therefore quite weak, as mentioned in Chapter 3. Another problem is that the two subjective judgments of Importance and Prominence were made contiguously and it is possible that a

blurring of definitions may have resulted in an undesirable influencing of responses (see Chapter 5).

In order to claim unitariness, we would require correlations close to 1.0 between each of the components. Since this condition is not met, we conclude that salience is not a unitary concept.

Investigation of the second research question provides a somewhat mixed result. We asked whether salience is a stable concept. In other words, is there an effect for subject proficiency level? The answer appears to be: It depends. If the task is to indicate how important an error is to correct, then the group to which a subject belongs does not make a significant difference. In that respect, the Birdsong and Kassen (1988) results are replicated. Where present results diverge from theirs is in findings for absolute (i.e. as opposed to relative) importance ratings. As mentioned in the discussion in Chapter 3, absolute ratings (which they term "severity of judgment") reflect which part of the scale a subject tended to use most. Birdsong and Kassen found that teachers, especially native French speakers, judged error importance more harshly (i.e. used the upper end of the scale more) than did students. In fact, the investigators found that the only groups which did not differ significantly from each other in severity of judgment were nonnative teachers and intermediates. The present study does not include native speaking French teachers; nevertheless, we might have expected, based on the

Birdsong and Kassen results, to find differences in absolute ratings between beginners and teachers, and perhaps even between beginners and intermediates. No significant differences were found, although a trend may be observed with regard to the relativized resumptive pronoun constructions.

How can these findings be explained? As mentioned in Chapter 3, it is possible that students gain clues about which errors instructors consider most important, either by their reactions to errors in class or by grading procedures. In this way, judgments of relative error importance (that is, of one error type compared to another) may be more likely to become "matched" to those of instructors. It is also possible to discuss importance in terms of communicative value, as we have noted earlier in this dissertation. It may be that, irrespective of length of exposure to a given language, people agree on which error types most disrupt communication. In this study, however, none of the errors are obstructive to comprehension, so it is unlikely that communicative value plays a role.

If students agree with teachers about the importance of errors, why don't they correct them? Birdsong and Kassen (1988) point out that a precondition for judging error importance is the ability to detect errors. It may be that learners are poor error detectors. We find significant differences between all groups on the detectability measure, with respect to both reaction time and accuracy. Thus it

appears that teachers are significantly better at error detection than are intermediates, who are in turn better than beginners. In other words, detectability, unlike importance, does not appear to be a stable concept.

Prominence is not stable either. Although it appears that subjects realized how relatively easy or difficult an error was for them to spot, teachers found errors most prominent and beginners found them least prominent, and these differences were significant.

It should be noted that the best way to determine the stability or instability of salience would be to do longitudinal studies (see Chapter 5); nevertheless, strong evidence exists that salience, in terms of detectability and prominence, is not a stable concept.

The answer to the third question, namely whether there is an effect for the type of error, is affirmative across the board. The problem is that the types of errors judged most important or most prominent are not necessarily associated with the fastest reaction times. For example, error type 3 (double object dative construction) was judged least important and least prominent, but is nevertheless associated with the fastest reaction times. On the other hand, error type 4 was judged most important and most prominent, and also received next to the fastest reaction times. Possible reasons for these results include slight differences in length of stimulus sentences and in location of errors. These issues are discussed in Chapter 5.

Error types 4 and 5, and, to a lesser degree, 6 (all of which are relativized resumptive pronoun constructions) are judged significantly more important and more prominent than the other error types. Yet types 5 and 6 are associated with the slowest reaction times. Recall that unlike the other three error types, all three of the relativized resumptive pronoun constructions are ungrammatical in English as well as in French. Also recall, however, that Birdsong et al. (1984) found that these error types were judged less severely by L2 learners than were others (see section 2.3.2).

Now that we have answered the three research questions, we will examine several frameworks to which a clarified notion of salience is relevant.

4.2. Frameworks to Which Salience Is Relevant

4.2.1. Bialystok and Ryan's Metacognitive Framework

As discussed in Chapter 2, Bialystok and Ryan (1985a, 1985b) note that the traditional view of metalinguistic performance as a unique, linguistic performance distinct from general metacognitive performance has been largely unproductive. Accordingly, they propose a metacognitive framework for the investigation of questions like who has better metalinguistic performance and at what point performance can be said to be metalinguistic in nature (see

section 2.2.3). They give the following explanation (1985a: 230-231):

> Our approach is to assume that the term metalinguistic applies not to a specific mental accomplishment but rather to a set of problems which share certain features. The theoretical issue, then, is to determine what cognitive skills underlie the solutions to this set of problems. . . . Finally, within this conceptualization, it is not necessary to delineate a precise set of problems that are distinctively metalinguistic. . . . Rather, these problems would more appropriately be construed as existing along a continuum reflecting varying reliance on the underlying skills specified in the cognitive framework.

The two skills they identify and posit as dimensions in the framework are associated with structuring knowledge and accessing knowledge: respectively, analyzed knowledge and cognitive control.

Originally, Bialystok (1978: 72-73) distinguished between "implicit" and "explicit" representations of linguistic knowledge, saying that the former type of knowledge is demonstrable, but not amenable to analysis, while the latter can be tranferred to other contexts, verbally justified, etc. More recently, Bialystok and Ryan (1985a, 1985b) follow Bialystok (1982; cited in Chaudron, 1983) in eliminating the implicit/explicit dichotomy in favor of an "unanalyzed" to "analyzed" dimension of knowledge.

Analyzed knowledge involves a distinction, albeit a non-dichotomous one, between knowledge that is intuitive, or implicit, and knowledge that is explicit (v. Reber, 1976; Reber and Lewis, 1977). The distinction can also be made in

terms of unconscious vs. conscious knowledge (see section 1.3.2 for a discussion of the role of consciousness in L2 acquisition). Bialystok and Ryan do not, however, view consciousness as a criterion for determining analyzed vs. unanalyzed knowledge, but rather as an outcome of more advanced levels of analyzed knowledge. Furthermore, as previously mentioned, they place the dimension of analyzed knowledge on a continuum rather than positing a dichotomy of unanalyzed vs. analyzed knowledge. Increasing ability to categorize and structure knowledge (i.e. to represent the structure of knowledge as well as its meaning) is reflected in increased values on the analyzed knowledge dimension.

Cognitive control involves "the selection and coordination of information, usually within time constraints" (1985a, p. 235). The easier the functions of control are for an individual, the less effortful is performance (i.e. the less control is required). This notion is consonant with information processing theories of automatic vs. controlled processing (Schneider and Shiffrin, 1977; Shiffrin and Schneider, 1977); however, terminology becomes confusing, since Bialystok and Ryan's "improved cognitive control" appears to mean "automatic". Nevertheless, their formulation is also consistent with theories of selective attention (Posner, 1980, 1982) and effortful processing (Kahneman and Treisman, 1984). It is also compatible with Tunmer and Herriman's (1984) suggestion that a child's acquisition of the different metalinguistic

abilities depends on metacognitive development, i.e. his/her ability to control intellectual processes.

Although neither analyzed knowledge nor cognitive control are quantified in an absolute sense, the values on these dimensions increase steadily, generating a two-dimensional Cartesian space. Figure 4.1 shows the framework.

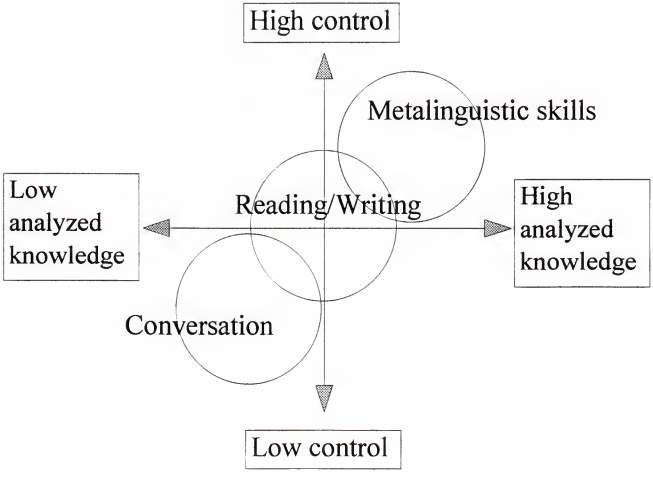


Figure 4.1. Cognitive framework underlying the development of language abilities. Adapted from Bialystok and Ryan, 1985.

As the figure shows, conversation, which requires minimal degrees of both analyzed knowledge and cognitive control, is located in the quadrant specifying low levels of both dimensions, while reading and writing, which require medium levels, are approximately centered over the four quadrants, and metalinguistic tasks, which require still higher levels of both skills, are in the upper quadrant. The space within each of the squares indicates that within each language domain, there is much variation in the demands along the two dimensions.

Bialystok and Ryan state that the dimensions of analyzed knowledge and cognitive control are theoretically independent, but related in two ways. First, a given task tends to demand relatively equal proportions of the two dimensions. Second, inadequate levels of one dimension constrain the expression of advanced levels of the other, since performance is an interaction of both. Bialystok and Ryan appear to suggest that this restriction is one of analyzed knowledge over control, but not vice-versa. In other words, degree of analyzed knowledge restricts expression of cognitive control; analysis of knowledge precedes control.

Menyuk (1985) states that there is a confusion in Bialystok and Ryan's articulation of the relationship between the dimensions. She argues that the claim that the dimensions are both sequential yet relatively independent is

contradictory, since ordering implies the existence of a dependency relationship between the two dimensions.

Bialystok and Ryan maintain that the relationship between the two dimensions is systematic, and not contradictory in terms of analyzing tasks. The confusion, they suggest, seems to arise from the developmental argument. They hypothesize that literacy is the primary contributor to the development of analyzed knowledge, while bilingualism is the major determinant of cognitive control.

It may be that there is an outcome of analyzed knowledge which is an additional promoter of cognitive control: an increase in the detectability and prominence of that knowledge (whether one or both are at issue will depend on whether the task is objective, subjective or both). Suppose that analysis of knowledge leads to increased detectability and prominence of that knowledge, which in turn results in improved cognitive control (i.e. less allocation of control resources) over the knowledge in question. Such a situation might help explain why analyzed knowledge precedes cognitive control of that knowledge. Note that it is not this researcher's intention to suggest that all knowledge over which a person has improved control is currently high in detectability and prominence, but only that at one time, the knowledge achieved levels sufficient to allow for easier control. Once knowledge is under cognitive control, it remains there, just as other automatic skills remain automatic. The fact that automatic processes

are difficult to alter or suppress has been noted by many investigators (Shiffrin and Schneider, 1977; Nation and McLaughlin, 1986, and others).

The notion that increased detectability and prominence may be an outcome of analyzed knowledge is not totally dissimilar to Bialystok and Ryan's suggestion that consciousness is an outcome of more advanced levels of analyzed knowledge. However, they do not, to the best of my knowledge, make the claim that consciousness affects control functions. Furthermore, they may not wish to make such a claim, since they do not view control as dependent on analyzed knowledge.

The present investigator suggests that there is a dependency relationship, however. In Chapter 1, we observed a circularity involved in current reasoning about salience: "What is salient gets noticed, and that which is noticeable is salient" (p.1). It may be that a more appropriate, and meaningful, way to view the situation is the following: What is analyzed becomes more detectable and prominent, and what is more detectable and prominent becomes more easily controlled (i.e. more automatic). Thus the degree to which a given construction has been analyzed determines the detectability and prominence of that construction, which in turn determines how easily controlled it is. Additionally, it is plausible to suggest that the frequency of a construction in the input contributes to analysis of that construction and that it is in this way (i.e. indirectly)

that frequency is a determinant of salience, at least when it is defined as detectability or prominence.

The purpose of this dissertation is neither to suggest a new model nor to revise an existing one; large amounts of data will be required before such an enterprise can be undertaken.

Salience can be seen to be relevant to two additional cognitive frameworks. One of these is Nelson's Cognitive Comparison Theory.

4.2.2. Nelson's Rare-Event Cognitive Comparison Theory

This theory has been formulated to explain how new syntactic constructions are introduced into a child's first language. However, Nelson (1981) suggests that aspects of second language acquisition may also fit this model. Tomasello and Herron (1989) suggest that the cognitive comparison process is the mechanism which underlies the effectiveness of their Garden Path technique (see section 1.3.3). According to Nelson's theory, the crucial learning experiences are those in which it is possible for learners to compare discrepancies between their own language structures and those of mature speakers. In his discussion of L1 acquisition, Nelson (1981: 231) states:

> The (typically) rare events absolutely necessary to the child's advances in syntax are the successful comparisons between input constructions and closely related constructions already in the child's syntactic system.

The reason the events are rare, Nelson says, is that discrepancies must be noted by the child. Limitations of memory, attention, and motivation all contribute to this rarity. He suggests that recasts (i.e. replies which retain the basic meaning of a child's sentence but utilize a new sentence structure) facilitate acquisition because they provide maximum opportunity for cognitive comparison.

Nelson stresses the necessity of noticing, on the part of the child. Similarly, in their L2 acquisition study, Tomasello and Herron found that their Garden Path technique, in which learners made hypotheses which were subsequently disconfirmed, constituted a more effective type of negative evidence. They hypothesized that the surprise factor increased the likelihood of a comparison.

Putting all of this together, it seems reasonable to suggest that a relatively more detectable and prominent linguistic construction is more likely to be "noticed", and thus subject to cognitive comparison, than is a relatively less detectable and prominent one. Since cognitive comparison is, according to the theory, crucial for learning, it follows that a high degree of detectability and prominence increases the likelihood that a construction will be learned. It is in this sense that salience can be seen to be relevant to the cognitive comparison framework. In the next section, we explore the other cognitive framework, that of connectionism.

4.2.3. The Connectionist Framework: PDP and the Competition Model

A recent trend in artificial intelligence studies of visual perception has involved approaches variously termed "parallel visual computation," "neo-associationism," "neo-connectionism" or "massive parallel processing systems" (Gardner, 1985). These approaches consider that it makes more sense to model visual processing using parallel systems (i.e. many systems carrying out many processes at the same time) rather than the traditional serial systems (i.e. systems operating in a sequential manner). Additionally, this view involves the notion that the independent processors not only carry out their own separate analyses, but exchange information with other processors in ways analogous to the simultaneous firing of brain cells or columns. Computation is effected by excitatory and inhibitory interactions among a large network of units. These simple neuronlike units compete and cooperate so that certain units become active and others are suppressed; in other words, their effects are achieved statistically. Information is not stored in any particular location; instead, it is widely distributed over the entire system (Rumelhart and McClelland, 1982, and others).

Two main advantages to this view can be discerned (Gardner, 1985): The first is that the existence of multiple connections allows knowledge from the entire system to be brought to bear on any given problem. The second

advantage is that even though many units (or processors) may be destroyed, widespread distribution of information ensures that destruction of the entire system does not occur. A related advantage is that it is possible to arrive at a decision even under conditions of noise.

Although the system mentioned above was originally proposed to deal with problems of visual processing, the observation that the system possesses characteristics which closely approximate those which an organism must logically be assumed to have in order to function in the chaotic natural world has not escaped the attention of researchers in other domains of cognitive science or even in other disciplines entirely. Parallel distributed processing (PDP) builds on developments such as those described above and, since its 1986 appearance as a major theory of cognition, has attracted interest in research domains as diverse as visual perception, motor control, memory, and language (Schmidt, 1988). We focus our attention here on its utility to theories of second language acquisition.

Traditionally, the dominant view in SLA theory, as a reflection of the influence of generative grammar, has been that learners possess a system of implicit (unconscious), abstract rules. The language acquisition process is conceived of as the setting of parameters for these rules on the basis of linguistic input and with the aid of an innate component that severely constrains the set of possible rules. In contrast, PDP theory claims that there exist no

rules as such, but only association strengths distributed across the entire network. Learning occurs as a result of the strengthening and weakening of network connections in response to examples encountered in the input. That it may occasionally appear as if rules are being learned is one interesting aspect of this type of learning.

Although Sampson (1987) is enthusiastic about the potential role of the PDP paradigm in linguistic theory and research, Schmidt (1988:63) points out that for many investigators, PDP and other connectionist models may recall unpleasant memories of behaviorism and thus be unappealing. He suggests that neither "enthusiastic conversion" nor "quick dismissal" are appropriate responses currently. He states that some PDP-inspired accounts of SLA should appear imminently and that if they do not, this will be a "disappointing sign of isolation from the intellectual mainstream". Other researchers believe that PDP models have much to offer the field of SLA (Sokolik, 1990; Spolsky, 1988). Larsen-Freeman (1991b) points out that any model which learns without access to rules and which can account for at least "some performance without postulating competence" (Spolsky, 1989: 227) may well end up causing us to "rethink earlier assumptions" (Larsen-Freeman, 1991b: 325).

Schmidt notes that the fact that the PDP model produces generalizations on the basis of whatever input is attended to or noticed by the learner makes it compatible with a view

of language learning in which consciousness and attention to form are required. We have seen in Chapter 1 that both consciousness and attention are fairly inextricably linked with salience, thus it seems fair to suggest that there may be a role for salience in the PDP paradigm.

Schmidt predicts that the most important role of PDP theory might be that of calling attention to connectionist models in general. One connectionist model which has received attention in second language acquisition research is the Competition Model (Bates and MacWhinney (1982). According to their model, there are two levels of language with which to communicate: form and function. One of the tasks involved in sentence comprehension is the assigning of function to sentence elements by means of surface form cues. This is not a simple task, since the mapping between form and function is many-to-many, rather than one-to-one. For example, a function such as the patient role (who or what the action of the verb is performed on) can be indicated by word order, morphological marking, etc. A form such as a morphological marker (e.g. a prefix, suffix, etc.) can indicate agent, patient, or a variety of other functions. Thus speakers use form cues to determine the function of elements in a sentence. Functions include agent, patient, beneficiary, instrument, etc. While languages do not all have the same functions represented, all languages use some subset of a common set of functions. Types of cues which aid interpretation include word order, morphological

markers, and animacy, among others, and these cues vary from language to language. In addition, the strength of cues varies. In other words, languages of the world differ not only in cue inventory but also in the strength of a given cue. Thus one of the findings of research in cue theory is that different languages rely on different cues for sentence comprehension (Bates and MacWhinney, 1982; Bates et al., 1982; Gass, 1989; McDonald and Heilenman, 1991a, 1991b).

The Competition Model, like most current cognitive models, assumes an existence of limited processing resources. In natural languages, relationships among elements in a sentence can be determined through examination of morphological markers, word order, lexical items, or prosody. Not all of these cues are equally used in a given language. In fact, sometimes they come into "competition" during processing; that is, cues can suggest conflicting interpretations of a sentence or an utterance. Resolution of conflicts is based on the strength of the cues involved. According to the model, cues in a given language are assigned a strength relative to the validity of the cue in that language. Validity is defined as the number of times that a cue is both present and indicates the correct role assignment. Otherwise stated, cue validity depends both on frequency and reliability of the cue in the language. A cue which is frequent in a language and usually correct will have a high strength and will defeat a cue that has lower validity (McDonald and Heilenman, 1991a).

The notion of cue validity originated in L1 acquisition research (see Bates et al., 1982, for one such study). Its impetus derives from its usefulness in explaining both learners' acquisition order of cues and the relative ranking of cues by adult speakers. However, its application has recently been broadened to include second language acquisition. Because of cross-linguistic differences in cues for sentence comprehension, one of the problems facing a second language learner is the acquisition of the correct mapping for the new language. Researchers have been interested in what happens when learners are faced with a language whose cue structure differs from that of their native language.

Research on second language learners within the Competition Model framework has found results which range from the transfer of L1 strategies, to the successful adoption of new strategies appropriate to the L2, with what appears to be a compromise in the middle, wherein learners adopt strategies intermediate to both languages (Bates and MacWhinney, 1981; Harrington, 1987; Gass, 1989; McDonald and Heilenman, 1991a, 1991b).

Native English speakers rely heavily on word order in sentence comprehension (Bates et al, 1982; McDonald and Heilenman, 1991a, 1991b). In contrast, native French speakers rely more heavily on (written) verb agreement and animacy than on word order (McDonald and Heilenman, 1991a). In a study of the acquisition of French by native English

speakers, McDonald and Heilenman (in press) found that learners changed their word order strategy before mastering the verb agreement cue. The investigators note that French verb endings are often not orally distinct, and suggest that the late increase in strength of the verb agreement cue may be a result of the "lack of general availability" of that cue. It will be recalled from Chap. 1 that Bardovi-Harlig (1987) defined salience, at least for her study, as "availability". Gass (1989:194) states that "noting the importance" of a cue such as word order is a necessary prerequisite to accurately "sorting out the details of the phenomenon". She suggests that this idea is consonant with a view of second language acquisition in which conscious awareness and selective attention are necessary conditions. This "sorting out the details of the phenomenon" would logically seem a prerequisite to an ability to deal with errors involving that phenomenon.[1]

McDonald and Heilenman (1991) state that recent research has indicated that cue validity (i.e. frequency and reliability of a cue) is not the only explanation of the order in which learners acquire cues and the relative ranking of cues by adult speakers. In addition to validity, "the order in which cues are acquired may be affected by the detectability or salience of a particular cue" (p.314).

It may be that "noting the importance" of a piece of input, whether it is a cue, a particular construction, or an error within a construction, is a prerequisite to being able

to detect (or find to be prominent) that input. Thus salience, in all three senses of the term, may play a role in this type of connectionist model.

4.3. Summary

In this chapter we have seen that increased detectability and prominence may be an output of analyzed knowledge and an input to cognitive control, within the metacognitive framework proposed by Bialystok and Ryan. Viewed from the perspective of Nelson's cognitive comparison framework, a high degree of detectability and prominence may increase the likelihood of cognitive comparison and thus maximize a structure's chances of being learned. Within connectionist frameworks such as the Competition Model, importance, prominence and detectability may all be determinants of the acquisition order of cues.

As previously mentioned, it is not the intention of this dissertation to present any formal modification of frameworks; such an enterprise would be premature. What is sought, however, is clarification of the nature of salience in second language acquisition. It is toward that end that the present studies were effected.

4.4. Notes

1. A logical extension of this idea is that noting the importance of a particular construction is a prerequisite to dealing with errors involving that construction. Yet in

section 4.1, we noted Birdsong and Kassen's suggestion that an ability to detect errors is a precondition for judging their importance. This apparent circularity is a result of the fact that it is impossible to consider an error without also considering the construction within which the error is located. It is necessary, however, to make a distinction between importance of a construction and importance of an error.

CHAPTER 5
CONCLUSION

5.1. Limitations of the Research

There are numerous limitations to the present research. The most obvious is that all data were collected under artificial conditions. Subjects dealt only with visual stimuli. Furthermore, they were asked to imagine a real-world situation (e.g. in directions for Importance judgments) but were required to respond in an artificial manner, especially as they were asked (albeit not explicitly) to respond to form rather than to meaning. Another point to be made is that all of the stimulus sentences were isolated out of context. Thus generalizability of the results obtained in this study to real-world learning situations is by no means certain.

Performance problems exist for each of the tasks. Subjects occasionally hit the wrong key on the error detection task; this fact was reported to the investigator. The error verification task was problematic in that subjects may have misunderstood it (see section 3.4.3); subject correction of errors appears to be more reliable. A possible confound between Prominence and Importance judgments, not only inherently but also as a result of their

having been performed contiguously, has already been discussed (see Chapter 4). The ceiling effect noted (again, see Chapter 4) in the GEFT scores may be partially the result of the use of a relatively small group of subjects; nevertheless, it poses a significant problem. Perhaps a different test should be used. Certainly, the use of a larger number of subjects is indicated.

In addition to the ecological validity issue and subjects' performance problems, limitations exist with regard to the stimuli. Although sentence length was controlled for as much as possible (see section 3.3.1.4), there were small differences. Similarly, location of error was controlled for as much as possible, but some errors occurred earlier in sentences than did others.

All of these difficulties point to the advisability of a conservative approach to interpreting the present findings.

5.2. Problems in Interpretation

Other problems exist which should be kept in mind as a caveat in interpreting the results of this study. Even apart from the possible confounds already mentioned, it is impossible to be sure that inter- and intra-subject definitions for each of the two subjective judgments of importance and prominence remained constant. In other words, even if a subject performed only one of the judgment

tasks, there is no guarantee that his or her definition was the same throughout the task, in spite of directions given.

As with any language-oriented experiment, particular exemplars chosen may influence the outcome. This point is made in discussion of what has been termed the "language-as-fixed-effect fallacy" (see Clark, 1973). It was with precisely this point in mind that analysis of the present data was carried out; for example, the three relativized resumptive pronoun construction subtypes were analyzed separately, rather than together. In spite of this precaution, more data will be needed before we can make definitive statements about the salience of types of errors (see section 3.3.3).

The novice/expert distinction is another potential problem for a study such as this one, since it is difficult to define "novice" and "expert" with any degree of exactitude. As mentioned in section 2.2.2, exposure and expertise do not necessarily correlate. For the purposes of the present study, the teacher group can be unambiguously classified as the "expert" group. Classification of the learners is potentially more problematic, except that in this case, nearly all of the learners, whether beginner or intermediate, have had at least a year of a foreign language other than French and received a grade of B or better in their last French grammar course (see section 3.3.2.4). Thus it appears as if the major distinguishing

characteristic between the learner groups in this study is length of exposure to French.

As mentioned in Chapter 4, the stability issue is most accurately examined by comparing performance of a given individual at time 1 with performance of that same individual at time 2. The present study is therefore an indirect investigation of stability, and caution must be employed in interpreting the results of the study.

5.3. Pedagogical Implications

One conclusion we arrive at as a result of this study is the same as that drawn by Birdsong and Kassen (1988): A "meeting of minds", with regard to error importance, seems to occur between teachers and students over time. Present results indicate that this is the case both for absolute ratings and for relative ratings.

It looks as if "input salience creation" of the kind discussed by Sharwood Smith (see section 1.2), where instructors and text book authors put visual flags in the input in hopes of creating mental flags in the minds of learners, is not a bad idea. However, "input salience augmentation" might be a more accurate term, and might result in improved control if used on the right construction types at the right time. The jury is still out on just what these "right" types and times are, however.

5.4. Directions for Further Research

We have discussed the possibility of a confound between Importance and Prominence judgments. One way to further investigate the nature of the association between these two subjective judgments is to conduct a follow-up study in which they are not both performed by all subjects, but rather each performed by different (but comparable) halves of the subject pool.

The question was posed in this dissertation as to whether salience is a stable concept, or whether what is salient changes with the learner's experience. Since what we are really asking is whether the salience of an error *to the same learner* changes, the best way to address the stability question, given all the individual differences which we have discussed, may be to do longitudinal studies on individual subjects.

5.5. Closing Remarks

The idea that salience is an important issue has not been sufficiently entertained in either the general literature or in the SLA literature. We have seen in previous chapters that a notion of salience has a contribution to make to many areas of SLA inquiry, such as theories of intake, consciousness and negative evidence, and to cognitive models such as Nelson's cognitive comparison, Bialystok and Ryan's Metacognitive framework, and cue theory.

In this dissertation, we have asked whether the concept of salience is a unitary one and whether it is stable. Additionally, we sought to discover whether some types of errors were more salient than others. Although results are somewhat mixed, it appears, so far, as if when we talk about a "salient" linguistic item in SLA, we need to specify salient *in what sense*, and salient *to whom*.

REFERENCES

Ackerman, P. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. Psychological Bulletin, 102 (1), 3-27.

Adelson, B. (1981). Problem solving and the development of abstract categories in programming languages. Cognition, 9, 422-433.

Adelson, B. (1984). When novices surpass experts: The difficulty of a task may increase with expertise. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 483-495.

Allwood, C. M. (1984). Error detection processes in statistical problem solving. Cognitive Science, 8, 413-437.

Anderson, J.R. (1985). Cognitive psychology and its implications. New York: W.H.Freeman.

Armstrong, S. L., L. R. Gleitman, and H. Gleitman. (1983). What some concepts might not be. Cognition, 13, 263-308.

Arthur, B. (1980). Gauging the boundaries of second language competence: A study of learner judgments. Language Learning, 30, 177-194.

Ashcraft, M.H. (1989). Human memory and cognition. Glenview, IL: Scott, Foresman and Company.

Ayoun, D. (1992). Parameter setting and learnability theory in the acquisition of French and Spanish by adult English native speakers. Doctoral dissertation, University of Florida.

Baars, B.J. (1988) A cognitive theory of consciousness. Cambridge: Cambridge University Press.

Baker, C.L., and J. McCarthy, eds. (1981) The Logical problem of language acquisition. Cambridge, MA.: MIT Press.

Bardovi-Harlig, K. (1987) Markedness and salience in second language acquisition. Language Learning, 37 (3), 385-407.

Barsalou, L.W. (1983). Ad hoc categories. Memory and Cognition, 11 (3), 211-227.

Barsalou, L. (1987). The instability of graded structures: Implications for the nature of concepts. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization. (pp. 101-140). Cambridge: Cambridge University Press.

Bates, E., McNew, S., MacWhinney, B., Devescovi, A. and Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. Cognition, 2, 245-299.

Bates, E., and MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner and L. R. Gleitman (Eds.), Language acquisition: The state of the art. Cambridge: Cambridge University Press.

Bates, E., and MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), Mechanisms of language learning. Hillsdale, NJ: Lawrence Erlbaum.

Ben Zeev, S. (1977). The influence of bilingualism on cognitive strategy and cognitive development. Child Development, 48, 1009-1018.

Bialystok, E. (1978). A Theoretical model of second language learning. Language Learning, 28, 69-83.

Bialystok, E. (1990). The dangers of dichotomy: A reply to Hulstijn. Applied Linguistics, 11 (1), 46-51.

Bialystok, E., and Ryan, E. B. (1985a). Toward a definition of metalinguistic skill. Merrill-Palmer Quarterly, 31, 229-264.

Bialystok, E., and Ryan, E. B. (1985b). A Metacognitive framework for the development of first and second language skills. In D. L. Forrest-Presley, G. E. MacKinnon, & T. G. Waller (Eds.), Meta-cognition, cognition, and human performance. New York: Academic Press.

Birdsong, D. (1989). Metalinguistic Performance and Interlinguistic Competence. Berlin: Springer-Verlag.

Birdsong, D. , C. Johnson, J. McMinn and P. Ingmundson (1984). Universals versus transfer revisited. Paper

presented at the Ninth Boston University Conference on Language Development.

Birdsong, D., and M. A. Kassen. (1988). Teachers' and students' evaluations of foreign language errors: A meeting of minds? Modern Language Journal, 72,(1), 1-12.

Birdsong, D., K. Thoren, and M.L. Price. (1984). "Variables affecting teachers' evaluations of errors in second language performance. Unpublished paper, University of Texas, Austin.

Bley-Vroman, R. (1986). Hypothesis testing in second language acquisition theory. Language Learning, 36, 353-376.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S.M. Gass and J. Schachter (Eds.), Linguistic perspectives on second language acquisition (pp. 41-68) Cambridge: Cambridge University Press.

Bley-Vroman, R.W., S.W. Felix, and G.I. Ioup. (1988) The accessibility of universal grammar in adult language learning. Second Language Research, 4, 1-32.

Boulouffe, J. (1986). Intake as the locus of equilibration in language learning. Language Learning, 36 (3), 245-275.

Brown, R. (1973). A first language. Boston: Harvard University Press.

Butler, B. (1980). Selective attention and stimulus localization in visual perception. Canadian Journal of Psychology, 34 (2), 119-133.

Carr, T. (1979). Consciousness in information processing. In G. Underwood and R. Stevens (Eds.), Aspects of Consciousness. London: Academic Press.

Cathcart, R. L., and J. E. W .B. Olsen. (1976). Teachers' and students' preferences for correction of classroom conversation errors. In F. Fanselow and R.H. Grymes (Eds.), On TESOL '76 (pp. 41-53). Washington, DC: TESOL.

Chafe, W. L. (1977). The recall and verbalization of past experience. In R.W. Cole (Ed.), Current issues in linguistic theory. Bloomington: IU Press.

Chase, W. G., and Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), Visual information processing. New York: Academic Press.

Chastain, K. (1980). Native speaker reaction to instructor-identified student second language errors. Modern Language Journal, 64, 210-215.

Chaudron, C. J. (1983a) The language of instruction in ESL classes: Variations in topic reinstatement and their effect on comprehension. Unpublished doctoral dissertation, University of Toronto, Canada.

Chaudron, C. (1983b). Research on metalinguistic judgments: A review of theory, methods, and results. Language Learning, 33, 343-377.

Chaudron, C. (1985a). A method for examining the input/intake distinction. In S. Gass and C. Madden (Eds.), Input in second language acquisition (pp. 285-300). Rowley, MA: Newbury House.

Chaudron, C. (1985b). Intake: On models and methods for discovering learners' processing of input. Studies in Second Language Acquisition, 7, 1-14.

Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge, MA: MIT Press.

Chomsky, N. (1981a). Markedness and core grammar. In A. Belleti, L. Brandi, and L. Rizzi (Eds.), Theory of markedness in generative grammar (pp. 123- 146). Pisa, Italy: Scuola Normale Superiore.

Chomsky, N. (1981b). Lectures on government and binding. Dordrecht, The Netherlands: Foris Publications.

Clahsen, H. (1990). The comparative study of first and second language development. Studies in Second Language Acquisition, 12, 135-153.

Clark, H. (1973) The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. Journal of Verbal Learning and Verbal Behavior, 12, 335-359.

Clark, H. H., and E. Clark. (1977). Psychology and language: An introduction to psycholinguistics. New York: Harcourt Brace.

Cook, V. (1985). Chomsky's universal grammar and second language learning. Applied Linguistics, 6, 2-18.

Cook, V. (1990). Timed comprehension of binding in advanced L2 learners of English. Language Learning, 40 (4), 557-599.

Cooper, W. E., and J. Paccia-Cooper. (1980). Syntax and speech. Cambridge, MA: Harvard University Press.

Coppieters, R. (1987). Competence differences between native and near- native speakers. Language 63, 544-573.

Corder, S.P. (1967). The significance of learners' errors. International Review of Applied Linguistics, 5, 161-170.

Corder, S.P. (1983) A role for the mother tongue. In S. Gass and L. Selinker, eds., Language transfer in language learning. Rowley, MA: Newbury House.

Crain, S. and J.D. Fodor. (1987). Sentence matching and overgeneration. Cognition, 26, 123-169.

Cunningham, J. G., and R. D. Odom. (1978). The role of perceptual salience in the development of analysis and synthesis processes. Child Development, 49, 815-823.

Delisle, H. (1982). Native speaker judgment and the evaluation of errors in German. Modern Language Journal, 66, 39-48.

den Heyer, K. (1986). Manipulating attention-induced priming in a lexical decision task by means of repeated prime-target presentations. Journal of Memory and Language 25, 19-42.

Deutsch, J., and D. Deutsch. (1963). Attention: Some theoretical considerations. Psychological Review, 70, 80-90.

Dulany, D.E., R.A. Carlson and G.I. Dewey (1984). A case of syntactical learning and judgement: How conscious and how abstract? Journal of Experimental Psychology: General, 113 (4), 541-555.

Dulany, D.E., R.A. Carlson, and G.I. Dewey. (1985). On consciousness in syntactic learning and judgment: A reply to Reber, Allen and Regan. Journal of Experimental Psychology: General, 114, 25-32.

Efron,R., Crandall, P. H., Koss, B., Divenyi, P. L., and Yund, E. W. (1983). Central auditory processing: III. The "cocktail party" effect and anterior temporal lobectomy. Brain and Language, 19, 254- 263.

Ellis, R. (1991). Grammaticality judgments and second language acquisition. Studies in Second Language Acquisition, 13, 161-186.

Ensz, K. (1982). French attitudes toward typical speech errors of American speakers of French. Modern Language Journal, 66, 133-139.

Ervin, G. (1978). A study of the use and acceptability of target language communication strategies employed by American students of Russian (Doctoral dissertation, The Ohio State University). Dissertation Abstracts International, 38, 6579A-6580A.

Fillmore, C.J. (1977). The case for case reopened. In P. Cole and J.M. Saddock (Eds.), Syntax and semantics: Vol 8. Grammatical relations. New York: Academic Press.

Galambos, S. J., and K. Hakuta. (1988). Subject-specific and task-specific characteristics of metalinguistic awareness in bilingual children. Applied Psycholinguistics, 9, 141-162.

Galloway, V. (1980). Perceptions of the communicative efforts of American students of Spanish. Modern Language Journal, 64, 429-433.

Gardner, H. (1985) The mind's new science. New York: Basic Books, Inc.

Gass, S. (1980). An investigation of language transfer in adult second language learners. In R. Scarcella and S. Krasher (Eds.), Research in second language acquisition (pp. 132-141). Rowley, MA: Newbury House.

Gass, S.M., and J. Ard. (1984). Second language acquisition and the ontology of language universals. In W. E. Rutherford, ed., Language universals and second language acquisition (pp. 33-68). Amsterdam: John Benjamins.

Gass, S.M., and J. Schachter (1989). Linguistic perspectives on second language acquisition. Cambridge: Cambridge University Press.

Glucksberg, S., and G.N. Cowen, Jr. (1970). Memory for nonattended auditory material. Cognitive Psychology, 1, 149-156.

Green, D. M., and Swets, J. (1966). Signal detection theory and psychophysics. New York: Wiley.

Gregg, K. (1986). Review of The input hypothesis: Issues and implications. TESOL Quarterly, 20, 116-122.

Gregg, K. R. (1989). Second language acquisition theory: the case for a generative perspective. In S.M. Gass and J. Schachter, (Eds.): Linguistic perspectives on second language acquisition (pp. 11-40). Cambridge: Cambridge University Press.

Harrington, M. (1987). Processing transfer: Language-specific processing strategies as a source of interlanguage variation. Applied Psycholinguistics, 8, 351-377.

Hatch, E. (1983). Psycholinguistics: A second language perspective. Rowley, MA: Newbury House.

Herron, C., and Tomasello, M. (1988). Learning grammatical structures in a foreign language: Modelling versus feedback. French Review, 61, 910-922.

Howard, D.V. (1983). Cognitive psychology: Memory, language and thought. New York: Macmillan.

Hulstijn, J. H. (1990). A comparison between the information-processing and the analysis/control approaches to language learning. Applied Linguistics, 11 (1), 30-45.

Ianco-Worrall, A. (1972). Bilingualism and cognitive development. Child Development, 43, 1390-1400.

Issidorides, D. C. (1988). The discovery of a miniature linguistic system: Function words and comprehension of an unfamiliar language. Journal of Psycholinguistic Research, 17 (4), 317-339.

Kahneman, D. and A. Treisman. (1984). Changing views of attention and automaticity. In R. Parasuraman and D. Daview (Eds.), Varieties of attention. New York: Academic Press.

Keenan, E., and B. Comrie. (1977). Noun phrase accessibility and universal grammar. Linguistic Inquiry, 8, 63-99.

Kellerman, E. (1983). Now you see it, now you don't. In S. M. Gass and L. Selinker, (Eds.), Language transfer in language learning. Rowley, Mass.: Newbury House.

Kellerman, E. (1985). If at first you do succeed... In S. Gass and C. Madden (Eds.), Input in second language acquisition (pp. 345-353). Rowley, MA: Newbury House.

Kellogg, R. (1980). Is conscious attention necessary for long term storage? Journal of Experimental Psychology: Human Learning and Memory, 6, 379-390.

Krashen, S.D. (1985). The input hypothesis: Issues and implications. London: Longmans.

Larsen-Freeman, D. (1978). An explanation for the morpheme accuracy order of learners of English as a second language. In E. Hatch (Ed.), Second language acquisition. Rowley, MA: Newbury House.

Larsen-Freeman, D. (1991). Second language acquisition research: Staking out the territory. TESOL Quarterly, 25 (2), 315-350.

Levelt, W. J. M. , J. A. W. M. Van Gent, A. F. J. Haans and A. J. A. Meijers. (1977). Grammaticality, paraphrase and imagery. In S. Greenbaum (Ed.), Acceptability in language. The Hague: Mouton.

Levelt, W.J.M. (1989). Speaking: From intention to articulation. Cambridge, MA: The MIT Press.

Lewicki, P., Hill, T., and Bizot, H. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. Cognitive Psychology, 20, 24-37.

Liceras, J. (1985). The role of intake in the determination of learners' competence. In S. Gass and C. Madden (Eds.), Input in second language acquisition. Rowley, MA: Newbury House.

MacWhinney, B. (1977). Starting points. Language, 53, 152-168.

Magnan, S. (1982). Native speaker reaction as a criterion for error correction. In A. Garfinkel (Ed.), ESL and the foreign language teacher (pp. 30-46). Skokie, IL: National Textbook.

Malt, B., and E. Smith. (1984). Correlated properties in natural categories. Journal of Verbal Learning and Verbal Behavior, 23, 250-269.

Maratsos, M. (1983). Some current issues in the study of the acquisition of grammar. In P.H. Mussen (Ed.), Handbook of Child Psychology: Vol 3. Cognitive development, ed. by J.H.Flavell and E.M. Markman (pp. 707-786). New York: Wiley.

Massaro, D. (1985). Attention and perception: An information-integration perspective. Acta Psychologica, 60, 211-243.

Mazurkewich, I. (1984). The acquisition of the dative alternation by second language learners and linguistic theory. Language Learning, 34 (1), 91- 109.

Mazurkewich, I. (1985). Syntactic markedness and language acquisition. Studies in Second Language Acquisition, 7 (1), 15-35.

McDonald, J. L., and L. K. Heilenman. (1991). Determinants of cue strength in adult first and second language speakers of French. Applied Psycholinguistics, 12, 313-348.

McLaughlin, B., T. Rossman, and D. McLeod. (1983). Second language learning: An information-processing perspective. Language Learning, 33, 135-158.

McLaughlin, B. (1987) Theories of second language learning. London: Edward Arnold.

McLaughlin, B. (1990). Restructuring. Applied Linguistics, 11 (2), 113-128.

Menyuk, P. (1985). Wherefore metalinguistic skills? A commentary on Bialystok and Ryan. Merrill-Palmer Quarterly, 31 (3), 253-259.

Nagata, H. (1988) The relativity of linguistic intuition: The effect of repetition on grammaticality judgments. Journal of Psycholinguistic Research, 17 (1), 1-17.

Naiman, N., M. Frolich, H.H. Stern, and A. Todesco. (1975). The good language learner. Toronto: Ontario Institute for Studies in Education.

Nation, R., and B. McLaughlin. (1986). Novices and experts: An information processing approach to the "good language learner" problem. Applied Psycholinguistics, 7 (1), 41-55.

Neisser, M. (1976). Cognition and reality. San Francisco, CA: W.H. Freeman.

Nesdale, A. R., and W. E. Tunmer. (1984). The development of metalinguistic awareness: A methodological overview. In W.E. Tunmer, C. Pratt., and M.L. Herriman (Eds.), Metalinguistic awareness in children: Theory, research and implications. Berlin: Springer-Verlag.

Nissen, M. J., and P. Bullemer (1987). Attentional requirements of learning: Evidence from performance measures. Cognitive Psychology, 19, 1-32.

Norman, D. (1968). Toward a theory of memory and attention. Psychological Review 75(6), 522-536.

Odlin, T. (1986). On the nature and use of explicit knowledge. IRAL, 24, 123-144.

Odom, R. D. (1982). Lane and Pearson's inattention to relevant information: A need for the theoretical specification of task information in developmental research. Merrill-Palmer Quarterly, 28 (3), 339-345.

Odom, R. D., E. C. Astor, and J. G. Cunningham. (1975). Effects of perceptual salience on the matrix task performance of four- and six-year-old children. Child Development, 46, 758-762.

Odom, R. D., and C. M. Lemond. (1975). The recall of relevant and incidental dimensional values as a function of perceptual salience, cognitive set, and age. Journal of Experimental Child Psychology 19, 524-535.

Oller, J. W. Jr. (1988). Review of S. Krashen, The Input Hypothesis: Issues and implications. Language, 64, 171-173.

O'Malley, J., A. Chamot, and C. Walker. (1987). Some applications of cognitive theory to second language acquisition. Studies in Second Language Acquisition, 9, 287-306.

Osgood, C.E., and J.K. Bock. (1977). Salience and sentencing: Some production principles. In S. Rosenberg (Ed.), Sentence production. Hillsdale, NJ: Lawrence Erlbaum.

Pachella, R. (1974). The interpretation of reaction time in information-processing research. In B. Kantowitz (Ed.), Human information processing: Tutorials in performance and cognition. Hillsdale, NJ: Lawrence Erlbaum.

Paivio, A., J.C. Yuille, and S.A. Madigan (1968). Concreteness, imagery and meaningfulness values for 925 nouns. Journal of Experimental Psychology (monograph supplement), 76, (1), Part 2.

Peters, A. (1985). Language segmentation: Operating principles for the perception and analysis of language. In D. Slobin, (Ed.), The cross-linguistic study of language acquisition, (Vol I.) Hillsdale, NJ: Erlbaum.

Piazza, L.G. (1980). French tolerance for grammatical erors made by Americans. The Modern Language Journal, 64 (4), 422-427.

Pienemann, M. (1984). Psychological constraints on the teachability of languages. Studies in Second Language Acquisition, 6 (2), 186-214.

Pinker, S. (1984). Language learnability and language development. Cambridge, MA: Harvard University Press.

Politzer, K. (1978). Errors of English speakers of German as perceived and evaluated by German natives. The Modern Language Journal, 62, 5-6, 253-261.

Polomska, M. (1988). A case for 'acquisitional strategies': Some methodo-logical observations on investigation into second language learners' initial state. Second Language Research, 4 (2), 110-132.

Posner, M. (1980). Orienting of attention. Quarterly Journal of Experimental Psychology, 32, 3-25.

Posner, M. (1982). Cumulative development of attentional theory. American Psychologist, 37 (2), 168-179.

Posner, M., and C. Snyder. (1980). Attention and cognitive control. In R. L. Solso (Ed.), Information-processing and cognition: The Loyola symposium. Hillsdale, NJ: Erlbaum.

Posner, M. , Snyder, C., and Davidson, B. (1980). Attention and the detection of signals. Journal of Experimental Psychology: General, 109, 160-174.

Reber, A.S. (1976). Implicit learning of synthetic languages: The role of instructional set. Journal of Experimental Psychology: Human Learning and Memory, 2, 88-94.

Reber, A.S., Kassin, S. M., Lewis, S., and Cantor, G. (1980). On the relationship between implicit and explicit modes of learning a complex rule structure. Journal of Experimental Psychology: Human Learning and Memory, 6, 492-502.

Reber, A.S., R. Allen, and S. Regan. (1985). Syntactical learning and judgment: Still unconscious and still abstract. Journal of Experimental Psychology: General, 114, 17-24.

Robb, T., S. Ross, and I. Shortreed. (1986). Salience of feedback on error and its effect on EFL writing quality. TESOL Quarterly, 20 (1), 83-93.

Rutherford, S. (1982). Markedness in second language acquisition. Language Learning, 32 (1), 85-107.

Searle, J.R. (1990). Consciousness, explanatory inversion, and cognitive science. Behavioral and Brain Sciences, 13, 585-642.

Scarcella, R., and L. Perkins. (1987). Shifting gears: Krashen's input hypothesis. Studies in Second Language Acquisition, 9, 347-354.

Schachter, J. (1983). A new account of language transfer. In S. M. Gass and L. Selinker (Eds.), Language transfer in language learning. Rowley, MA: Newbury House.

Schachter, J. (1986). Three approaches to the study of input. Language Learning, 36, 211-225.

Schachter, J. (1989). Testing a proposed universal. In S. M. Gass and J. Schachter (Eds), Linguistic perspectives on second language acquisition (pp. 73-88). Cambridge, MA: Cambridge University Press.

Schmidt, R. W. (1988). The potential of parallel distributed processing for SLA theory and research. University of Hawai'i Working Papers in ESL, 17 (1), 55-66.

Schmidt, R. W. (1990). The role of consciousness in second language learning. Applied Linguistics, 11 (2), 129-158.

Schneider, W. and R. M. Shiffrin. (1977). Controlled and automatic human information processing: I. Detection, search and attention. Psychological Review, 84 (1), 1-54.

Schneider, W., S. T. Dumais, and R. M. Shiffrin. (1984). Automatic and control processing and attention. In R. Parasuraman and D.R. Davies (Eds.), Varieties of attention (pp. 1-27). Orlando, FL: Academic Press.

Schneiderman, E. I. and C. Desmarais. (1988). The talented language learner: Some preliminary findings. Second Language Research, 4 (2), 91-109.

Scholes, R. J. and B. J. Willis. (1987). Language and literacy. Journal of Literary Semantics, 16, 3-11.

Scribner, S., and M. Cole. (1981). The psychology of literacy. Cambridge, MA: Harvard University Press.

Sharwood Smith, M. (1986). The competence/control model, crosslinguistic influence and the creation of new grammars. In E. Kellerman and M. Sharwood Smith (Eds.), Crosslinguistic influence in second language acquisition (pp. 10-20). New York: Pergamon Press.

Sharwood Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. Second Language Research, 7 (2), 118-132.

Shiffren, R. M., and W. Schneider. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.

Snow, C., and G. Meijer. (1977) On the secondary nature of syntactic intuitions. In S. Greenbaum (Ed.), Acceptability in language. The Hague: Mouton.

Sperling, G. (1984). A unified theory of attention and signal detection. In R. Parasuraman and D.R. Davies (Eds.), Varieties of attention. New York: Academic Press.

Strickland, K., and D. Ayoun (1989). Relative salience of sentential elements for second semester French students. Unpublished paper, University of Florida, Gainesville.

Thoren, K. (1988). Selective attention and its application to theories of L2 intake. Unpublished paper, University of Florida, Gainesville.

Tomasello, M., and C. Herron. (1988) Down the garden path: Inducing and correcting overgeneralization errors in the foreign language classroom. Applied Psycholinguistics, 9, 237-246.

Tomasello, M., and C. Herron. (1989). Feedback for language transfer errors. Studies in Second Language Acquisition, 11, 385-395.

Treisman, A. M., and G. Gelade. (1980). A feature-integration theory of attention. Cognitive Psychology, 12, 97-136.

VanPatten, W.C. (1984) Processing strategies in second language acquisition Unpublished doctoral dissertation: University of Texas, Austin.

VanPatten, B. (1989). Can learners attend to form and content while processing input? Hispania, 72, 409-417.

Velmans, M. (1989) Is human information processing conscious? Paper, University of London (draft).

White, L. (1985). Is there a logical problem of second language acquisition? TESL Canada, 2 (2), 29-41.

White, L. (1987a). Markedness and second language acquisition: The question of transfer. Studies in Second Language Acquisition, 9, 261-286.

White, L. (1987b). Against comprehensible input: The input hypothesis and the development of L2 competence. Applied Linguistics, 8, 95-110.

White, L. (1989). The adjacency condition on case assignment: do L2 learners observe the subset principle? In S.M. Gass and J. Schachter (Eds.), Linguistic perspectives on second language acquisition. Cambridge: Cambridge University Press.

White, L. (1990). Second language acquisition and universal grammar. Studies in Second Language Acquisition, 12, 121-133.

Winograd, P., and P. Johnston. (1982) Comprehension monitoring and the error detection paradigm. Journal of Reading Behavior, 14 (1), 61-75.

Witkin, H. A., Oltman, P. K., Raskin, E., and Karp, S. A. (1971). A manual for the embedded figures tests. Palo Alto, CA: Consulting Psychologists Press, Inc.

APPENDIX 1
QUESTIONNAIRE

experiment ID # _____

Please do not provide your name. Information on this questionnaire is strictly confidential.

1. Have you ever studied or been exposed in any way to a language other than English or French (including even a brief exposure long ago)? YES NO

If you answered "YES" to this question, please answer questions 2-4. If you said "NO", you may skip 2-4.

2. To what language(s) were you exposed?
_____
3. How many years of exposure would you estimate you have had? ____ yrs.
4. How many years has it been since you were last exposed? _____ yrs.

All subjects should respond to the remaining questions.

5. How difficult have you found learning French to be, on a scale of 1 (not at all difficult) to 6 (very difficult)?
1 2 3 4 5 6

6. What was the last grammar-oriented French course you took? (In other words, if you are an advanced student, do not count literature courses.)
_____
7. What grade did you receive in your last grammar course?
_____
8. How much outside-class study time would you estimate you averaged per week during that course? _____ hours per week
9. What grades have you received in any other French courses taken? (Include only college courses, from any university or college.)

Course	_____	Grade	_____
Course	_____	Grade	_____
Course	_____	Grade	_____
Course	_____	Grade	_____

10. Compare yourself to other students in your last class. If the best students ranked "1" and the weakest students "6", where do you think you would be on the scale in terms of:

a) speaking proficiency	1	2	3	4	5	6
b) knowledge of French	1	2	3	4	5	6

APPENDIX 2
STIMULI

Preposition Stranding

Questions (error type 1)

1. *Qui est-ce que Luc voit le film avec?
2. *Qui est-ce que Jean achète le livre pour?
3. *Qui est-ce que Nathalie vote contre?
4. *Qui est-ce que Charles part sans?
5. *Qui est-ce que Dianne met le chapeau sur?

Relative clause (type 2)

1. *Voilà le stylo que Robert écrit avec.
2. *Voilà l'examen que Suzanne étudie pour.
3. *Voilà l'ami que Claire compte sur.
4. *Voilà l'appartement que Thomas habite dans.
5. *Voilà le mur que Mireille place l'armoire contre.

Double object dative construction (type 3)

1. *Marie achète Nathalie une robe.
2. *Paul donne Suzanne un livre.
3. *Françoise écrit Jean une lettre.
4. *Bernard envoie Nadine un chèque.
5. *Lucile prête Henri un stylo.

Relativized resumptive constructions

Subject (type 4)

1. *La femme qui elle est à la porte est sympa.
2. *La fille qui elle est très belle parle allemand.
3. *La femme qui elle est française est riche.
4. *La femme qui elle est dans la rue est blonde.
5. *La fille qui elle est célèbre porte des lunettes.

Direct Object (type 5)

1. *La femme que Jean la voit à la porte est sympa.
2. *La fille que Marcel la regarde toujours est très belle.
3. *La femme que Luc la déteste est riche.
4. *La femme que Richard la cherche dans la rue est blonde.
5. *La fille que Charles la préfère est célèbre.

Indirect Object (type 6)

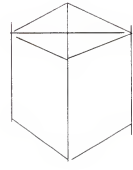
1. *La femme que Jean lui parle est sympa.
2. *La fille que Marcel lui téléphone toujours est très belle.
3. *La femme que Luc lui donne son adresse est riche.
4. *La femme que Richard lui demande l'heure est blonde.
5. *La fille que Charles lui vend un billet est célèbre.

* Indicates ungrammaticality

APPENDIX 3
SAMPLE GEFT ITEM

Subjects are instructed to find the simple (triangle) form within the more complex form. Simple forms are located on the back of the test booklet; they cannot be viewed at the same time as the complex forms. Directions:

> Try to find the simple form in the complex figure and trace it in pencil directly over the lines of the complex figure. It is the same size, in the same proportions, and faces in the same direction within the complex figure as when it appeared alone.



APPENDIX 4
EXPERIMENT 2 RAW DATA

Subject	Percent Correct	Mean RT	GEFT score
01	76.9	05.0 sec	14
02	67.5	08.2 sec	18
03	97.5	03.4 sec	17
04	65	11.6 sec	10
05	75	02.3 sec	15
06	50	15.3 sec	18
07	70	16.5 sec	16
08	87.5	14.9 sec	18
09	84.6	03.2 sec	15
10	95	10.5 sec	18
11	56.7	04.3 sec	18
12	95	12.2 sec	17
13	87.5	05.2 sec	18
14	84.6	11.2 sec	16
15	80	04.4 sec	09
16	76.9	13.4 sec	08
17	58.9	03.1 sec	17
18	62.5	05.6 sec	08

BIOGRAPHICAL SKETCH

Krista Helene Thoren was born in Los Angeles and grew up in Bloomington, Indiana. She received her bachelor's degree in French from Indiana University in 1983. She earned a master's degree in French linguistics from the University of Texas in 1985.

She has enjoyed teaching, not only at U.T. and U.F. but also at the Université Paul Valéry in France.

In May 1993 she married Carl Turner. They hope to do a lot of travelling.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

David Birdsong, Chairman
Associate Professor of Linguistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Andrea Tyler
Assistant Professor of Linguistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jean Casagrande
Professor of Linguistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ira Fischler
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

M. Jeffrey Farrar
Associate Professor of Psychology

This dissertation was submitted to the Graduate Faculty of the Program in Linguistics in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May, 1993

_____

Dean, Graduate School